

ЦИФРОВАЯ ОБРАБОТКА СИГНАЛОВ

Научно-технический журнал
№ 3/2017

Издается с 1999 года
Выходит четыре раза в год

ГЛАВНЫЙ РЕДАКТОР
д.т.н., чл. кор. РАН Ю.Б. ЗУБАРЕВ

ЗАМЕСТИТЕЛИ ГЛАВНОГО РЕДАКТОРА:

д.т.н., проф. В.В. ВИТЯЗЕВ,
д.т.н., проф. В.П. ДВОРКОВИЧ

РЕДАКЦИОННАЯ КОЛЛЕГИЯ:

д.т.н., проф. Ар.С. Аджемов, д.т.н., проф. Б.А. Алпатов,
д.т.н., проф. В.Г. Бартенев, д.т.н. Ю.И. Борисов,
д.т.н., проф. Ю.А. Брюханов, д.т.н., проф. В.В. Еремеев,
д.т.н. Г.В. Зайцев, Р.В. Зубарев, А.П. Кирпичников,
д.т.н., проф. М.И. Кривошеев, д.т.н., акад. РАН Н.А. Кузнецов,
В.Г. Мистюков, д.т.н., проф. С.Л. Мишенков,
д.т.н., проф. А.А. Петровский, д.т.н., проф. Ю.Н. Прохоров,
д.т.н., проф. В.Г. Санников, к.т.н., проф. В.С. Сперанский,
д.т.н., проф. Ю.С. Шинаков

Адрес редакции:

г. Москва, ул. Авиамоторная, д. 8
Научный центр МТУСИ
Тел.: (+7) 903-201-53-33
E-mail: rntores@mail.ru
vityazev.v.v@rsreu.ru
http://www.dspa.ru

Издатель:

Российское научно-техническое общество
радиотехники, электроники и связи им. А.С. Попова
Компьютерная верстка: И.А. Благодарова
Дизайн: М.В. Аверин

Подписной индекс по каталогу
ОАО «Роспечать» – **82185**

Подписано в печать 17.10.17 г.
Формат 60x90/8.

Гарнитура «Arial». Печать офсетная.
Бумага офсетная. Печ.л. 6,5. Тираж 500 экз.

Заказ № 6363. Отпечатано в
ООО НПЦ «Информационные технологии»
Рязань, ул. Островского, д. 21/1
тел.: (4912) 98-69-84

Издание зарегистрировано в Министерстве
Российской Федерации по делам печати,
телерадиовещания и средств
массовых коммуникаций.

Свидетельство о регистрации ПИ № 77-1488
от 14.01.2000 г.

Журнал «Цифровая обработка сигналов» включен
в Перечень ведущих рецензируемых научных изданий,
в которых по рекомендации Минобрнауки РФ, должны
быть опубликованы основные научные результаты дис-
сертаций на соискание ученой степени доктора
и кандидата наук.

УЧРЕДИТЕЛИ: ЗАО «Инструментальные системы» ФГУП «НИИ радио»

В НОМЕРЕ:

- Дворкович А.В., Дворкович В.П., Седова М.А.**
*«Цифровая экономика», мировая практика разработки
высокоэффективных цифровых систем телерадиовещания
и проблемы их внедрения в России* 3
- Боков А.А., Ватолин Д.С.**
*Восстановление фона в видео
с использованием непараметрической модели движения
и покадрового уточнения*..... 13
- Новиков А.И.**
*Решение переопределенных систем
линейных алгебраических уравнений
в задачах обработки изображений*..... 22
- Третьяков В.А., Кротков А.Ю.,
Кривошеев В.В., Данилов Р.Ю.**
*Улучшение процесса тематической обработки
гиперспектральной информации* 28
- Шипко В.В.**
*Алгоритм комплексирования двух разнесспектральных
изображений на основе цветового синтеза*..... 33
- Кузнецов А.Е., Рыжиков А.С.**
*Исследование быстродействия алгоритмов поиска
на изображениях одноименных точек
с использованием современной вычислительной технике*..... 40
- Хрящев В.В., Приоров А.Л.,
Стефаниди А.Ф., Топников А.И.**
*Разработка и исследование алгоритмов обработки
и распознавания речевых сигналов и изображений
для систем мультимодальной биометрии*..... 45
- Зайцев А.В., Соловьев В.А.,
Кичулкин Д.А., Красавцев О.О.**
*Сравнительная оценка способов обработки цифровых
изображений, получаемых с помощью тепловизора
при наблюдении беспилотных летательных аппаратов*..... 50
- Корепанов С.Е., Смирнов С.А., Стротов В.В.**
*Алгоритм определения дальности до объекта
для стереоскопической системы технического зрения
реального времени*..... 57
- Ершов М.Д., Шубин Н.Ю.**
*Алгоритмы обработки изображений
для решения задач анализа дорожной обстановки* 63
- Носков А.А., Волохов В.А.,
Приоров А.Л., Апальков И.В.**
*Синтез полностью сфокусированных изображений
при наличии шумов*..... 68
- Сорока Е.З.**
*К вопросу о влиянии светочувствительной матрицы
на качество изображения* 74
- Кононов В.С.**
*Разработка архитектуры КМОП-АЦП
с широким диапазоном частот преобразования
от нескольких килогерц до нескольких гигагерц*..... 78

Подписной индекс по каталогу
ОАО «Роспечать» – **82185**

Digital Signal Processing

Science & Technical Magazine
Issue 3, 2017 year

Is published quarterly since 1999

THE EDITOR-IN-CHIEF:

Dr.Sci. (Tech.), Professor, Corresponding Member
of Russian Academy of Sciences U.N. Zubarev

DEPUTY EDITORS-IN-CHIEF:

Dr.Sci.(Tech.), Professor V.V. Vityazev, Dr.Sci.(Tech.),
Professor V.P. Dvorkovich

EDITORIAL BOARD:

Adzhemov A.S., Dr.Sci.(Tech.), Professor

Alpatov B.A., Dr.Sci.(Tech.), Professor

Bartenev V.G., Dr.Sci.(Tech.), Professor

Borisov Y.I., Dr.Sci.(Tech.)

Bruchanov Y.A., Dr.Sci.(Tech.), Professor

Eremeyev V.V., Dr.Sci.(Tech.), Professor

Zaitsev G.V., Dr.Sci.(Tech.)

Zubarev R.V.,

Kirpichnikov A.P.,

Krivocheev M.I., Dr.Sci.(Tech.), Professor

Kuznetsov N.A., Dr.Sci.(Tech.), Academician

Mistyukov V.G.,

Mishenkov S.L., Dr.Sci.(Tech.), Professor

Petrovsky A.A., Dr.Sci.(Tech.), Professor

Prokhorov Y.N., Dr.Sci.(Tech.), Professor

Sannikov V.G., Dr.Sci.(Tech.), Professor

Speranskii V.S., Dr.Sci.(Tech.), Professor

Shinakov Y.S., Dr.Sci.(Tech.), Professor

Editorial office address:

Aviamotornaya, 8, Moscow, Russia

Research Center of MTUCI

Phone: (+7) 903-201-53-33

E-mail: rntores@mail.ru,

vityazev.v.v@rsreu.ru

Web: <http://www.dspsa.ru>

Publisher:

Russian A.S. Popov Society for Radioengineering,
Electronics & Communications

Computer makeup:

I.A. Blagodarova

FOUNDERS:

InSys

Radio Research and Development Institute

CONTENTS:

Dvorkovich A.V., Dvorkovich V.P., Sedova M.A.

«Digital Economy», worldwide practice of the development
of highly efficient digital TV and radio broadcasting systems
and the problems of its implementation in Russia3

Bokov A.A., Vatolin D.S.

Video background reconstruction by non-parametric motion
completion and frame-by-frame refinement13

Novikov A.I.

The solution of overdetermined systems of linear equations
in problems of image processing22

Tretiakov V.A., Krotkov A.Y.,

Krivoshein V.V., Danilov R.Y.

Improving the thematic processing of hyperspectral information28

Shipko V.V.

The algorithm of integrating two images
of different spectrums based on color synthesis33

Kuznetsov A.E., Ryzhikov A.S.

The analysis of the speed of search algorithms on images
of the corresponding points using modern computer40

Khryashchev V.V., Priorov A.L.,

Stefanidi A.F., Topnikov A.I.

Development and analysis of speech
and image signals processing
for multimodal biometrics systems45

Zaitsev A.V., Solovjov V.A.,

Kichulkin D. A., Krasavtsev O.O.

Comparative estimation of ways of processing digital images,
received by means of thermal imager under observation
of unmanned aerial vehicle50

Korepanov S.E., Smirnov S.A., Strotov V.V.

Algorithm for determining the distance to object
for real-time stereoscopic technical vision system57

Ershov M.D., Shubin N.Yu.

Image processing algorithms
for road situation analysis63

Noskov A.A., Volokhov V.A.,

Priorov A.L., Apalkov I.V.

Full-focused image fusion in the conditions of noise68

Soroka E.Z.

About impact of photo sensor matrix on image quality74

Kononov V.S.

Design of architecture of CMOS-ADC with wide bandwidth
from few kilohertz to few gigahertz78

Subscription index: 82185 ("Rospechat")

ISSN: 1684-2634 9 771684 263005

УДК 004.421, 004.932

«ЦИФРОВАЯ ЭКОНОМИКА», МИРОВАЯ ПРАКТИКА РАЗРАБОТКИ ВЫСОКОЭФФЕКТИВНЫХ ЦИФРОВЫХ СИСТЕМ ТЕЛЕРАДИОВЕЩАНИЯ И ПРОБЛЕМЫ ИХ ВНЕДРЕНИЯ В РОССИИ

Дворкович А.В., д.т.н., чл.-корр. РАН, директор физтех-школы радиотехники и компьютерных технологий МФТИ, e-mail: dvork-alex@yandex.ru;

Дворкович В.П., д.т.н., профессор, заведующий кафедрой мультимедийных технологий и телекоммуникаций МФТИ, e-mail: v.dvorkovich@mail.ru;

Седова М.А., директор по развитию и внедрению системы РАВИС, e-mail: sedovamarina2017@yandex.ru.

«DIGITAL ECONOMY», WORLDWIDE PRACTICE OF THE DEVELOPMENT OF HIGHLY EFFICIENT DIGITAL TV AND RADIO BROADCASTING SYSTEMS AND THE PROBLEMS OF ITS IMPLEMENTATION IN RUSSIA

Dvorkovich A.V., Dvorkovich V.P., Sedova M.A.

Digital information technologies under wide use and implementation should provide the high efficiency of «digital economy» utilization, that actualizes the modern trends of TV and radio broadcasting and communication. The variants of highly efficient digital TV and sound broadcasting and videoconferencing development based on the latest advances in multimedia coding and transmission technology have been analyzed. The results of the new broadcasting systems development allow to state positively about the possibilities of principal change of the information presentation quality. Broadcasting systems don't realize such functional possibilities up to now. Taking into account these world achievements, it should be mentioned that they are not deployed in Russia. Insufficient attention is paid to the implementation of digital TV technologies of standard and high definition, deployment of national broadcasting systems that provide sharp increase of radio spectrum usage efficiency, etc.

Key words: digital economy, digital TV and radio broadcasting, videoconferencing, multimedia encoding, quality of service QoS, quality of experience QoE.

Ключевые слова: цифровая экономика, цифровое телерадиовещание, видеоконференцсвязь, кодирование мультимедиа, качество сервиса QoS, качество восприятия QoE.

Введение

Стратегическая важность реализации «Цифровой экономики» в России отражена в Указе Президента РФ Путина В.В. № 203 от 09.05.2017 [1] и в Программе «Цифровая экономика Российской Федерации», утвержденной распоряжением Правительства Российской Федерации № 1632-р от 28.07.2017 [2].

Широко применяемые и внедряемые в настоящее время технологии используют широкополосные ресурсы гигагерцового диапазона частот и могут применяться только в больших городах, поскольку не обеспечивают качественное функционирование системы на огромных территориях РФ.

Из Указа Президента РФ Путина В.В. № 203 необходимо выделить следующие основные аспекты, имеющие непосредственное отношение к реализации «Цифровой экономики»:

– «Принять меры поддержки **традиционных** средств распространения информации, и в первую очередь – **радио** на всей территории РФ».

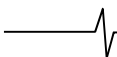
– Все Интернет-технологии таковыми не являются. Радио всегда было для большинства населения тради-

Широко применяемые и внедряемые цифровые информационные технологии должны обеспечить высокую эффективность использования «цифровой экономики», реализующей современные тенденции радиотелевизионного вещания и связи. Проведен анализ работ по развитию вариантов высокоэффективных систем цифрового телевизионного и звукового радиовещания и видеоконференцсвязи, разработки которых основаны на использовании последних мировых достижений технологий кодирования и передачи мультимедийного контента. Результаты создания новых вариантов систем телерадиовещания позволяют с уверенностью сказать о возможностях принципиального изменения качества воспроизводимой информации. До настоящего времени эти системы не обеспечивали таких функциональных возможностей. Учитывая эти новые мировые достижения, следует отметить, что они не реализуются в России, где недостаточно внимания уделяется внедрению даже цифровых технологий телевидения стандартной и высокой четкости, реализации отечественных разработок телерадиовещания, обеспечивающих резкое повышение эффективности использования частотного пространства, и др.

ционным, любимым и бесплатным сервисом. Интернет-технологии и технологии мобильной связи для доставки мультимедийного контента таковыми не являются.

– «Обеспечить насыщение рынка доступными, качественными и легальными сервисами российского производства».

– «**Ключевым направлением** повышения конкурентоспособности отечественных информационных техно-



логий является **радиотехника** и внедрение отечественных информационных технологий, формирование представления о внедрении инноваций как о **приоритетном пути технологического развития**».

Указ № 203 подчеркивает необходимость: «...развивать различные образовательные технологии, в том числе дистанционные, при реализации образовательных программ», «создать различные технологические платформы для дистанционного обучения в целях повышения доступности качественных образовательных услуг».

Кроме того, рекомендуется «стимулировать фундаментальные и прикладные научные исследования в сфере информационных технологий, выполняемых научно-техническими организациями, а также разработку высокотехнологического оборудования в этой сфере».

Указ Президента РФ Путина В.В. № 203 требует «осуществить интеграцию российских стандартов в сфере информационных технологий в соответствующие международные стандарты».

Непосредственное отношение к проблемам реализации «Цифровой экономики» имеет также Указ Президента РФ Путина В.В. № 208 от 13.05.2017 [3], который указывает на одну из основных угроз экономической безопасности – «стремление развитых государств использовать свои преимущества в уровне развития экономики, высоких технологий, в том числе информационных, в качестве инструмента глобальной конкуренции».

Внедрение отечественных информационных технологий должно создать противодействие этим угрозам и защищать национальные интересы, поскольку имеет место «слабая инновационная активность, отставание в области разработки и внедрения новых и перспективных технологий, в том числе технологий цифровой экономики в условиях, когда уже существующие Интернет-технологии не работают или неэффективны технически или экономически, например, в удаленных и труднодоступных районах».

Мы гордимся работами выдающихся ученых России и Советского Союза по созданию систем радио и телевидения. Всех этих уникальных ученых не перечислить, но всем специалистам известны имена А.С. Попова, Б.Л. Розинга, И.Л. Зворыкина, Г.В. Брауде, С.В. Новаковского, С.И. Катаева, П.В. Шмакова и многих, многих других [4, 5].

С сожалением можно утверждать, что в последние годы участие отечественных ученых в развитии цифрового телерадиовещания следует считать незначительным.

Во всем мире активно ведутся работы по развитию новых вариантов систем телевизионного вещания, обеспечивающих повышение качества передаваемой информации и зрительского восприятия. К этим работам относятся технологии телевидения ультравысокой четкости (ТУВЧ), расширенного динамического диапазона (HDR), повышенной скорости передачи и воспроизведения кадров видеoinформации (HFR), многоканального звукового сопровождения с эффектом погружения и пр.

Такие технологии обеспечивают зрителю более широкий диапазон цветов и яркостей, более плавные движения деталей изображения без мельканий, более

естественный звук и т.д., создавая эффект вовлечения зрителя в демонстрируемую сцену.

Выбранные на диаграмме МКО точки основных цветов определяют цветовую модель RGB: $R_1G_1B_1$ и белый цвет C определяют исходную цветность системы приемника стандарта NTSC; $R_2G_2B_2$ и белый цвет D_{65} соответствуют исходной цветности систем приемников стандартов PAL и SECAM; $R_3G_3B_3$ и белый цвет D_{65} характеризуют цветности приемников HDTV (рис. 1).

Каждый элемент (пиксел), представленный с помощью трех основных цветов RGB, определяется 24 битами (по 8 бит на красную, зеленую и синюю составляющие) для стандартного ТВ или 36 битами для HDTV/UHDTV (4K или 8K).

В табл. 1 даны координаты основных цветов систем цветного и цифрового телевидения [6].

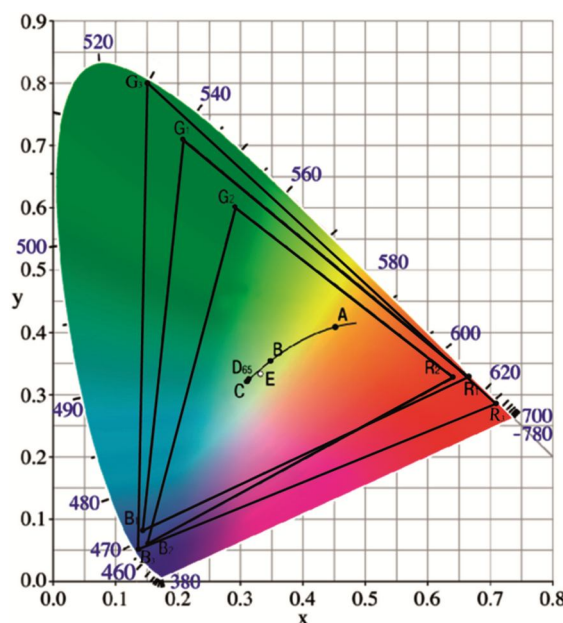


Рис. 1. Таблица цветностей МКО

Таблица 1. Координаты RGB цветов аналоговых и цифровых систем телевидения

Основные цвета	Системы телевизионного вещания					
	NTSC		PAL/SECAM		UHDTV	
	x	y	x	y	x	y
Красный R	0,67	0,33	0,64	0,33	0,708	0,292
Зеленый G	0,21	0,71	0,29	0,60	0,170	0,797
Синий B	0,14	0,08	0,15	0,06	0,131	0,046
Белый	0,310	0,316	0,313	0,329	0,3127	0,3290

На рис. 2 приведены варианты используемых яркостей воспроизводимых изображений (под рисунком приведены типы некоторых дисплеев) [7]. Как следует из этого рисунка, дисплеи стандартного ТВ обеспечивают изменение яркостей от 0,1 до 100 кд/м². Последние же реализации дисплеев (Dolby Research YDR Display) увеличивают этот диапазон в сотни раз (от 0,005 до 20 000).

На рис. 3 приведены варианты многоканального звукового сопровождения телевидения [8].

Ниже приведена таблица отключений аналогового вещания в Европе за 2015 год, где указано, что практически все страны выполнили эту задачу, за исключением Украины [9]. Только Россия определена пустой графой в табл. 2.

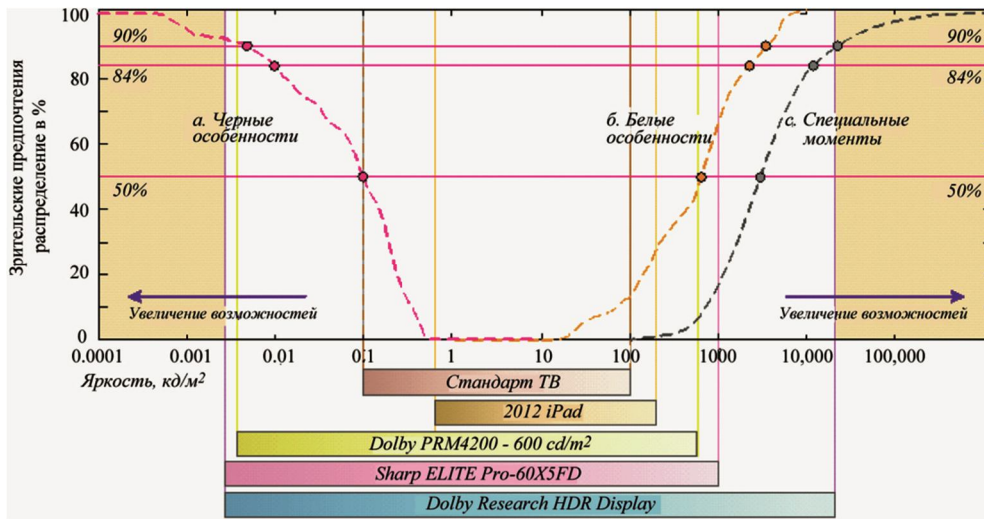


Рис. 2. Варианты используемых яркостей ТВ изображений (от стандартного ТВ до HDR)

Звуковая система	Высший слой	Средний слой	Нижний слой
H(9+10+3)	TrFL, TrFC, TrFR, TrBL, TrBC, TrBR	FLc, FC, FRc, EL, EC, ER	LFE1, LFE2, BFL, BR
Высший слой 3/3/3 (ТОН)	TrFL, TrFC, TrFR	FLc, FC, FRc	LFE1, LFE2
Средний слой 5/2/3	TrBL, TrBC, TrBR	EL, EC, ER	BFL, BR
Нижний слой 3/0/0.2	TrBC, TrBR	EC, ER	BFL, BR
G(4+9+0)	Lc, C, Rc	Lc, C, Rc	LFE
Высший слой 2/0/2	Lc, C, Rc	Lc, C, Rc	LFE
Средний слой 5/2/2	Lc, C, Rc	Lc, C, Rc	LFE
Нижний слой 0.0/0.1	Lc, C, Rc	Lc, C, Rc	LFE
F(3+7+0)	LH, RH	Lc, C, Rc, LH, RH	LFE1, LFE2
Высший слой 2/0/1	LH, RH	Lc, C, Rc	LFE1, LFE2
Средний слой 3/2/2	LH, RH	Lc, C, Rc, LH, RH	LFE1, LFE2
Нижний слой 0/0/0.2	LH, RH	Lc, C, Rc, LH, RH	LFE1, LFE2
V(0+5+0)		Lc, C, Rc	LFE
Высший слой 0/0/0		Lc, C, Rc	LFE
Средний слой 3/0/2		Lc, C, Rc	LFE
Нижний слой 0/0/0.1		Lc, C, Rc	LFE

Рис. 3. Различные стандарты звукового сопровождения телевидения

После 17 июня 2015 года не обеспечивается защита от помех для службы аналогового ТВ вещания соседних стран, что создает существенные сложности для ТВ вещания по всей протяженности границ России.

В табл. 3 приведен процент охвата населения цифровым телевидением в некоторых странах [10]. В табл. 4 приводятся ориентировочные годы завершения аналогового телевизионного вещания в ряде стран (приведенная для России дата была бы успехом полного перехода в России на цифровое телевидение) [10].

Следует заметить, что в России уже принят 51 государственный стандарт по цифровому вещанию (1 – в 2004 г., 3 – в 2012 г., 6 – в 2011 г., по 7 – в 2015 и 2016 г.г., по 8 – в 2011 и 2013 г.г. и 11 – 2014 г.), однако, в основном, они ориентированы на документы ISO, ITU, ETSI, их качество оставляет желать лучшего.

Некоторые ГОСТы предполагают использование стандартов, не предусмотренных для внедрения в Рос-

сии (например, DVB-H). В то же самое время не проводятся мероприятия по внедрению рассмотренных выше новейших технологий. По этой причине не следует удивляться тому, что аналоговое телевидение в России существует до сих пор.

Таблица 2. Даты официального перехода на цифровой формат в Европе

Страна	Дата запуска	Формат компрессии	Факт завершения
Великобритания	1998	MPEG-2	Завершено
Швеция	1999	MPEG-2	Завершено
Испания	2000/2005	MPEG-2	Завершено
Финляндия	2001	MPEG-2	Завершено
Швейцария	2001	MPEG-2	Завершено
Германия	2002	MPEG-2	Завершено
Бельгия	2002	MPEG-2	Завершено
Нидерланды	2003	MPEG-2	Завершено
Италия	2004	MPEG-2	Завершено
Франция	2005	MPEG-2 / MPEG-4 AVC	Завершено
Чехия	2005	MPEG-2	Завершено
Дания	2006	MPEG-2 /MPEG-4 AVC	Завершено
Эстония	2006	MPEG-4 AVC	Завершено
Австрия	2006	MPEG-2	Завершено
Словения	2006	MPEG-4 AVC	Завершено
Норвегия	2007	MPEG-4 AVC	Завершено
Литва	2008	MPEG-4 AVC	Завершено
Венгрия	2008	MPEG-4 AVC	Завершено
Украина	2008	MPEG-4 AVC	2014
Латвия	2009	MPEG-4 AVC	Завершено
Португалия	2009	MPEG-4 AVC	Завершено
Хорватия	2009	MPEG-2	Завершено
Польша	2009	MPEG-4 AVC	Завершено
Словакия	2009	MPEG-2	Завершено
Ирландия	2011	MPEG-4 AVC	Завершено
Россия	2012	MPEG-4 AVC	

Изменения систем телевизионного вещания связаны с реализацией лучших проектов цифрового телевидения. Так, в этом году в Германии запланирована так называемая «вторая цифровая трансформация» – переход от вещания DVB-T с кодированием видео по стандарту H.264 к DVB-T2 с кодированием видео H.265.

Таблица 3. Процент охвата населения цифровым телевидением

Страна	Процент охвата населения
Германия	90 %
Гонконг	99 %
Япония	95 %
Кения	58 %
Малайзия	98 %
Намибия	70 %
Сенегал	90 %
ЮАР	88 %
Испания	98 %
Танзания	25 %
Таиланд	95 %
Великобритания	98,5 %

Таблица 4. Даты перехода на цифровое телевидение в ряде стран (включая Россию)

Страна	Дата перехода
Австралия	2013
Финляндия	2007
Франция	2011
Германия	2008
Гана	2017
Индонезия	2018
Италия	2012
Япония	2011
Кения	2015
Малайзия	2020
Маврикий	2013
Намибия	2015
Нидерланды	2006
Нигерия	2016
Польша	2013
Россия	2018
Руанда	2014
ЮАР	2017
Южная Корея	2012
Испания	2010
Тайвань	2012
Танзания	2014
Таиланд	2020
Уганда	2015
Великобритания	2012
США	2009
Вьетнам	2020

В Китае внедряется принятый в 2006 году стандарт DTMB [11]. Разработан усовершенствованный вариант стандарта DTMB-A [12], несовместимый с DTMB. Его внедрение пока не запланировано.

Стандарт DTMB-A использует несколько передовых технологий канального кодирования и кадровой структуры, которые обеспечивают быструю синхронизацию, высокую чувствительность приема, повышенную устойчивость при многолучевом приеме, высокую спектральную эффективность и гибкость для будущих расширений стандарта. Полоса OFDM сигнала составляет 7,56 МГц в стандартном режиме и 70/9 МГц в расширенном режиме.

Система DTMB-A поддерживает двухантенную разнесенную передачу. Доступные скорости передачи данных варьируются в диапазоне от 5 до 50 Мбит/с для полосы 7,56 МГц.

Лабораторные и полевые испытания подтвердили превосходную производительность системы как при

фиксированном, так и при мобильном приеме в сложных условиях передачи.

Решение некоторых проблем эффективной обработки изображений HDTV в России изложено в энциклопедических трудах [13, 14], дополненных теоретическими исследованиями по применению многоканальных и комбинированных вейвлет-фильтров [15].

Для повышения эффективности преобразования изображений предложено производить разделение сигналов при помощи трехполосных, четырехполосных или пятиполосных фильтров и последующим дополнительным вейвлет-преобразованием отдельных высокочастотных областей.

В этом случае вейвлет-преобразование выполняется по стандартной схеме: сначала преобразованию подвергаются все строки изображения, в результате чего формируются 3-5 областей, затем такое же преобразование применяется ко всем столбцам полученных коэффициентов, в результате чего создается 9-25 областей. Затем производятся дополнительные преобразования отдельных высокочастотных областей с другими вейвлет-фильтрами [15].

На рис. 4 приведено преобразование изображения на 25 областей с использованием пятиполосного вейвлет-фильтра (рис. 5).

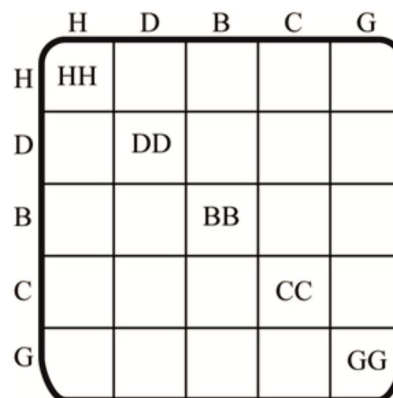


Рис. 4. Разбиение изображения на 25 поддиапазонов

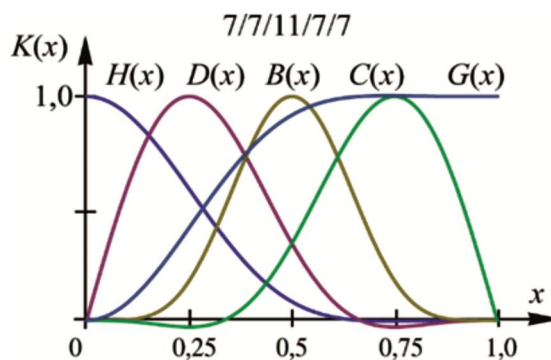


Рис. 5. Пятиполосный вейвлет-фильтр

Если учесть, что объем преобразованного в цифровой формат изображения сверхвысокой четкости составляет порядка 70 Гбит, то каждая область преобразованного изображения исходно имеет в среднем объем менее 3 Гбит, а дискретное преобразование и квантование каждой из 25 областей позволяет дополнительно уменьшить эту информацию в среднем не менее, чем в 100 раз. Следовательно, объем цифровой информации

каждой из областей составит в среднем порядка 30 Мбит. К тому же следует заметить, что такие преобразования могут эффективно осуществляться многопроцессорными системами.

На рис. 6 приведено изображение, преобразованное комплексными вейвлет-фильтрами на 16 поддиапазнов с применением трехполосного вейвлет-фильтра (рис. 7) и последующим разбиением ряда областей стандартным двухполосным вейвлет-фильтром.

HH		BHH		GHH
		BHG		GHG
HVH	HVG	BBHH	BBGH	GV
		BBHG	BBGG	
HGH	HGG	VG		GG

Рис. 6. Разбиение изображения на 16 областей

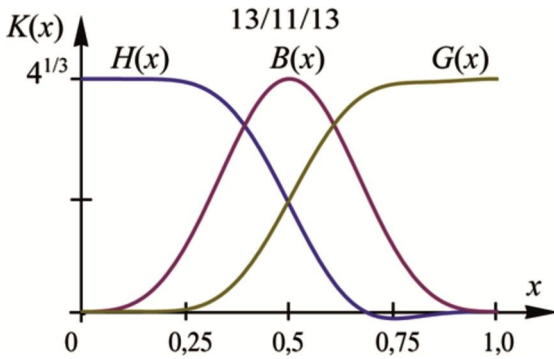


Рис. 7. Трехполосный вейвлет-фильтр

На рис. 8 в качестве примера приведено преобразованное на 16 поддиапазнов эталонное изображение «Залив». Анализируя это преобразованное изображение, можно заметить, что в среднечастотных и высокочастотных областях остается малый объем информации.

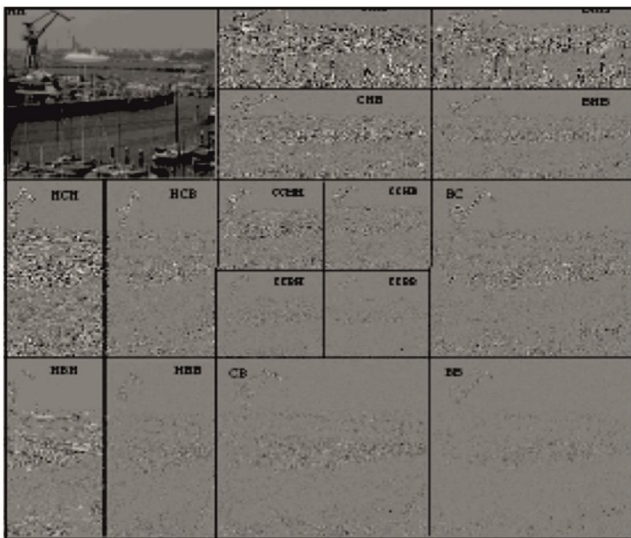


Рис. 8. Разбиение изображения «Залив» с применением комбинированных вейвлет-преобразований

На рис. 9 и 10 приводятся варианты обработки динамических изображений с использованием трехполосного трехуровневого преобразования поддиапазона HH. В данном случае изображение низкочастотной составляющей уменьшается по сравнению с исходным изображением в $9 \times 9 = 729$ раз.

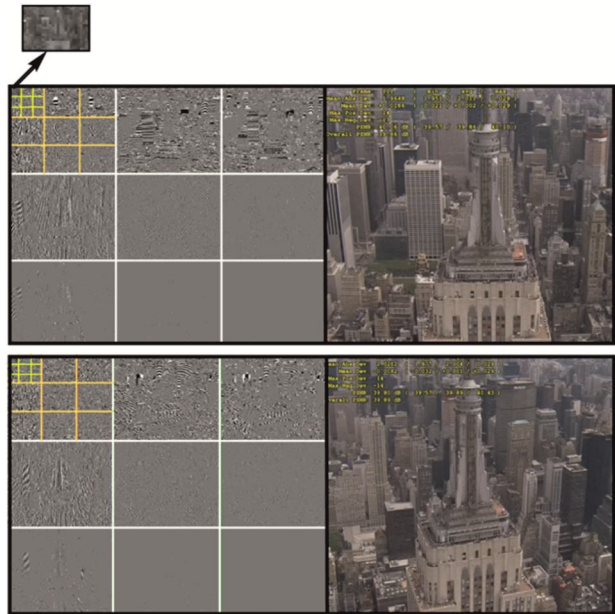


Рис. 9. Динамическое изображение «Город»

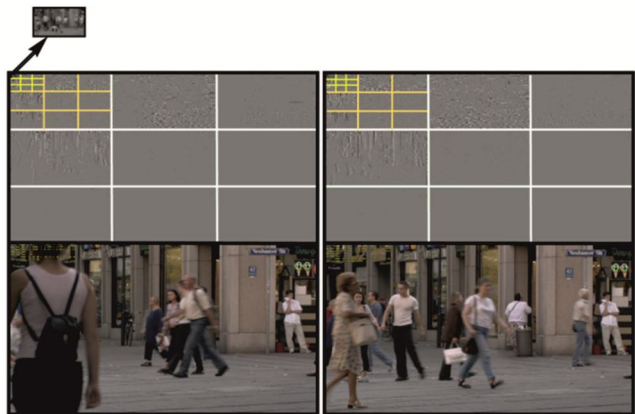


Рис. 10. Динамическое изображение «Движение в городе»

Постепенно происходит и внедрение цифрового звукового вещания, которое в настоящее время предусматривает формирование и передачу в I-III ОВЧ диапазонах мультимедийной информации, содержащей кроме трансляции стереофонического звукового вещания различного качества также передачу статических и динамических видеосигналов и других дополнительных цифровых данных.

В табл. 5 приведены основные характеристики четырех стандартов цифрового радио, внедряемых в ряде стран [16].

В системе DAB в едином мультиплексе (1,536 МГц) передается 6 основных программ, а также дополнительные данные, связанные с программами. Стереофонический звук MPEG Layer II со скоростью потока 192-256 кбит/с обеспечивает высокое качество, а со скоростью 128 кбит/с – нормальное качество. Услуги данных – текст, рисунки, web, радиогид, сообщения, информация о потоке. На рынке имеется более 350 моделей приемников, начиная от 25 \$.

Таблица 5. Основные характеристики систем цифрового радио

Характеристики	DAB/T-DMB/AT-DMB	HD Radio	DRM+	РАВИС
Полосы частот (МГц)	VHF III (174-230)	VHF II (87,5-108)	VHF I-III (47-230)	VHF I, II (66-74, 87,5-108)
Ширина канала	1,712 МГц	400 кГц	100 кГц	250/200/100 кГц
Передача данных	0,576-1,728 Мбит/с DMB: 0,864-2,304 Мбит/с	147-294 кбит/с	37-186 кбит/с	196-888/155-703/75-341 кбит/с
Спектральная эффективность, бит/с/Гц	DAB/DMB: 0,38-1,13 AT-DMB: 0,56-1,88	0,55-4,1	0,39-1,9	0,77-3,64
Данные, тип сервиса	Аудио, данные T-DMB, AT-DMB: видео	Аудио, данные	Аудио, данные	Видео, аудио, данные
Количество несущих	192/384/768/1536	25-27/38-191	213	553/439/215
Длительность сегмента	156/312/623/1246 мкс	5,504/2,752 мс	2,25 мс	2,25 мс
Метод модуляции	4-DQPSK	BPSK, QPSK, 16-QAM, 64-QAM	4-QAM, 16-QAM	4-QAM, 16-QAM, 64-QAM
Канальное кодирование	Сверточный код (1/4-3/4)+ +турбокод (1/4-1/2), RS(204,188,8)	Сверточный код (2/11-5/6), RS(255,191,32)	Сверточный мате- ринский код 1/6, 7 уровней	LDPC + BCH 1/2, 2/3, 3/4

DAB+ использует метод компрессии звука HE-AAC v2, мультиплекс включает от 12 до 18 каналов; скорость потока 64 кбит/с обеспечивает стереозвук хорошего качества, 48 кбит/с – нормального качества. Имеются несколько типов приемников и доступные чипсеты. Этот вариант стандарта внедряется в ряде стран в Европе и в Австралии. Норвегия полностью перешла на цифру в 2017 г.

T-DMB – мультимедийное радиовещание, использует стандарт H.264/AVC для кодирования видео и BSAC или HE-AAC v2 для звука. Обеспечивается передача 4 каналов со скоростью до 256 кбит/с. Имеется около 150 моделей приемников. Главным образом используется в Южной Корее.

HD Radio – метод передачи цифровых радиовещательных сигналов одновременно с традиционным радиовещательным сигналом (технология IBOC, 397,2 кГц). Активно применяется в США и Канаде. В ОВЧ диапазоне работает несколько тысяч радиостанций, все новые автомобили оснащаются приемниками, принимающими HDRadio. Фирма iBiquity (владелец торговой марки HD Radio и технологии IBOC) предпринимает большие усилия по распространению своей системы в мире много лет, но пока их усилия тщетны.

DRM+ – это расширение стандарта DRM30 на ОВЧ-диапазоны частот. DRM+ использует логическую структуру стандарта DRM, изменены только параметры OFDM модуляции для ширины полосы 96 кГц. Способ кодирования аудиоинформации – xHE-AAC.

DRM+ поддерживает скорость передачи данных от 40 до 186 кбит/с. В мультиплексе могут передаваться до 4 программ, например, возможна передача 2 музыкальных программ со скоростью 64 кбит/с и передача двух речевых программ или дополнительных данных со скоростью 24 кбит/с.

В формате DRM+ при модуляции 16-QAM мощность передатчика цифрового сигнала может быть на 12 дБ меньше мощности передатчика ЧМ сигнала при одинаковой зоне охвата. При QPSK модуляции разница мощности может составлять до 20 дБ. Регулярное вещание DRM+ не ведется.

Эксперименты по вещанию DRM+ проводились во многих странах мира, например, в Великобритании, Франции, Италии, Бразилии, Индии и др. В настоящее время имеется несколько вариантов коммерческих приемников.

РАВИС – единственное инновационное предложение России в телекоммуникационной сфере.

Принцип создания отечественной системы цифрового телерадиовещания РАВИС был обоснован патентами РФ на телерадиовещание через узкополосный канал связи [17, 18].

Основные преимущества системы РАВИС отражены в отечественных стандартах [19-23] и документах международных организаций [16, 24, 25], характеризующих международное признание системы. Система РАВИС обеспечивает:

- повышение эффективности использования диапазона ОВЧ (более чем в 10 раз);
- повышение качества звукового вещания, введение новых мультимедийных сервисов, в том числе видео;
- организацию системы локального и общероссийского оповещения;
- существенное снижение энергопотребления эфирного вещания;
- создание единой системы мобильного вещания при сохранении частотных распределений (не реализуется ни одной ТВ системой);
- организацию местного телевидения или радиовещания в малых городах и поселках, возможность работы без мультиплекса;
- реализацию высокоэффективных систем специального применения;
- интегрирование России в общемировую систему цифровых телекоммуникаций на базе российского стандарта и взаимодействие с другими стандартами радиовещания.

В сравнении с действующими системами цифрового радиовещания DAB/T-DMB/AT-DMB, HD Radio и DRM+, система РАВИС строго соответствует установленным в России частотным характеристикам и имеет существенные как технические, так и экономические преимущества.

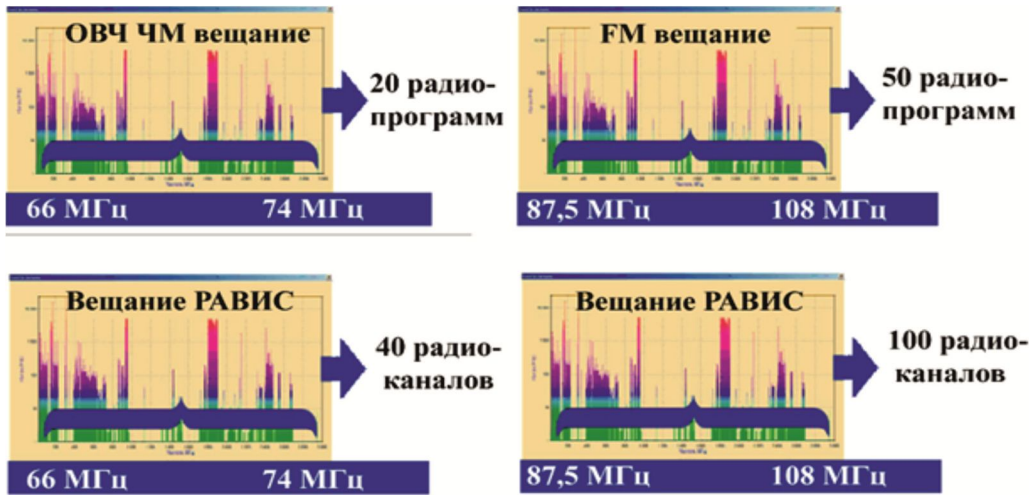


Рис. 11. Спектральная эффективность применения системы RAVIS

Рис. 11. иллюстрирует повышение эффективности использования частотного спектра. В одном радиоканале системы RAVIS может передаваться 10-20 стереофонических звуковых программ, либо до 8 многоканальных (5.1) звуковых программ, либо видеопрограмма и до 6 звуковых стереопрограмм.

На рис. 12 и 13 приведены примеры реализации одночастотных сетей в городе и вдоль железных и шоссейных дорог.

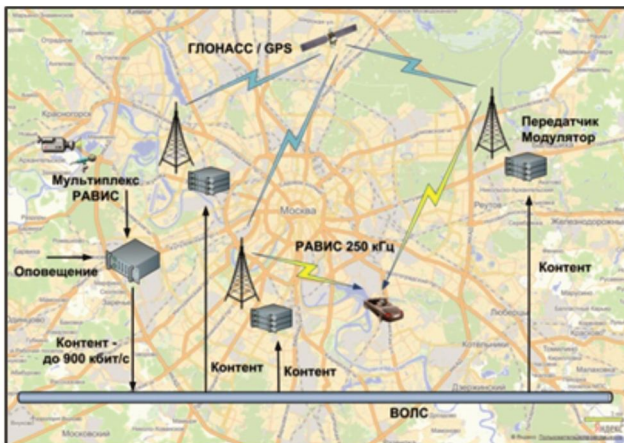


Рис. 12. Пример одночастотной сети с доставкой контента по кабельной сети

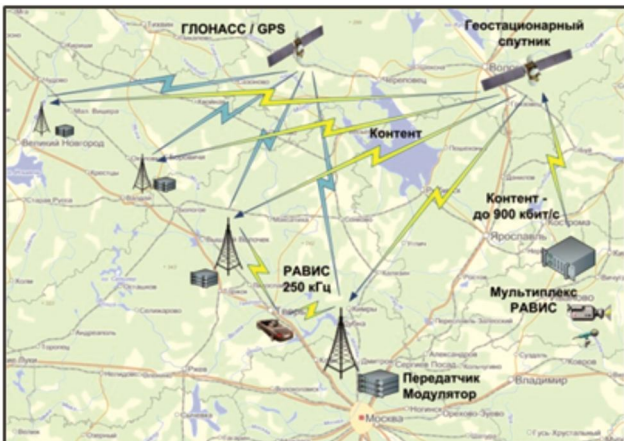


Рис. 13. Пример одночастотной сети с доставкой контента через спутник

Рис. 14 и 15 иллюстрируют работу системы RAVIS при фиксированном и мобильном варианте приема.

Разработка модели отечественной цифровой системы проводилась в 2003-2005 г.г. из средств Российского фонда технологического развития (РФТР).

В результате, осенью 2005 г. появились первые публикации в журналах, а в феврале 2006 г. – первые вклады России в МСЭ, после чего, по просьбе председателя консорциума DRM летом 2006 г. был сделан доклад на Техническом комитете DRM. Эти события совпадают по времени с началом активной разработки «конкурирующей» системы DRM+, более узкополосной и менее эффективной.

Работа по развитию системы RAVIS была положительно оценена в газете Россия № 1 от 17 января 2008 г.

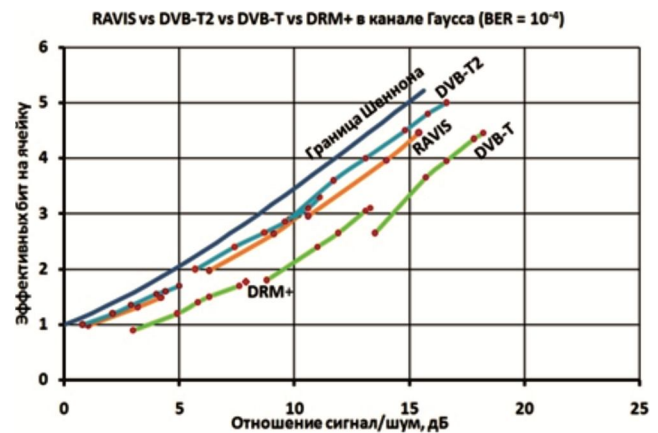


Рис. 14. Сравнение систем RAVIS, DVB-T, DVB-T2 и DRM+

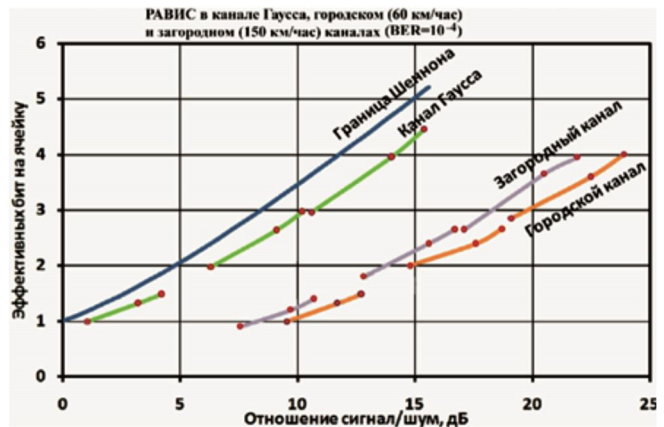


Рис. 15. Эффективность системы RAVIS при мобильном приеме

К сожалению, в то время «Министерство информатизации и связи России... не обеспечивало финансирование внедрения тех или иных разработок...».

Далее продолжение НИР осуществлялось в 2008-2009 г.г. по контракту с Федеральным Агентством по науке и инновациям. Работа была высоко оценена Заказчиком, было предложено подать документы для последующей реализации в рамках НИОКР. Однако, реализация этого предложения оказалась невозможна из-за ликвидации этого Федерального Агентства.

Международный симпозиум по цифровому радиовещанию (октябрь 2009 г.) принял Меморандум об объединении внедрения RABIS и DRM+, но консорциум DRM в дальнейшем от этих планов отказался.

19 августа 2009 г. ГКРЧ выделила частоты для организации тестовых испытаний системы RABIS, и такие испытания были проведены в Москве и Сочи. Результаты показали высокую эффективность системы, и в 2010 г., по докладу о результатах испытаний, ГКРЧ приняла рекомендацию Минпромторгу РФ включить в перечень ОКР разработку опытных образцов аппаратуры RABIS: «Разработка базовой технологии создания унифицированных модулей». Выполнение работ осуществлялось разработчиками RABIS в ГРЦ и ОАО «НПК НИИДАР».

Однако во внедрении системы RABIS Минкомсвязи было не заинтересовано, разработчикам системы в то время было сказано: «наука не нужна!», и хотя цифровая система радиовещания «имеет некоторые преимущества», но «оптимальным на сегодняшний день следует считать сохранение системы аналогового радиовещания как основы бизнес-модели».

Тем не менее, благодаря Никифорову Н.А., министру информатизации и связи, 16 октября 2015 г. ГКРЧ приняла решение о создании опытных зон для тестирования системы RABIS в городах Ижевске, Казани, Калининграде и Краснодаре.

За подписью министра Никифорова Н.А. были направлены письма руководителям регионов с предложением о запуске двух программ: передачи ТВ в городской транспорт и многопрограммной передачи звука мобильному потребителю. Было согласовано ТЗ и требуемый объем финансирования – около 10 млн. руб. на каждый город. Но необходимые средства до сих пор так и не выделены.

Это приводит к тому, что Россия лишается уникальных возможностей, которые создало бы для рынка внедрение технологии RABIS:

- развитие рынка вещания для мобильных абонентов;
- выход на рынок компонентов для мультимедийных автомобильных приемников;
- выход на рынок компонентов для бытовой аудио-видео аппаратуры;
- выход на рынок компонентов для смартфонов;
- выход на рынок компонентов для GPS/ГЛОНАСС навигаторов;
- выход на рынок дорожного информирования о пробках, дорожной ситуации, рекламы на дорожных цифровых картах;
- развитие рынка систем оповещения МЧС;
- развитие рынка систем специального применения;

– развитие отечественного производства и создание новых рабочих мест.

Одной из важных задач внедрения видеотелекоммуникационных систем в России является создание их метрологического обеспечения [26-28].

На рис. 16 приведена фотография одного из разработанных измерительных комплексов цифрового вещательного телевидения.

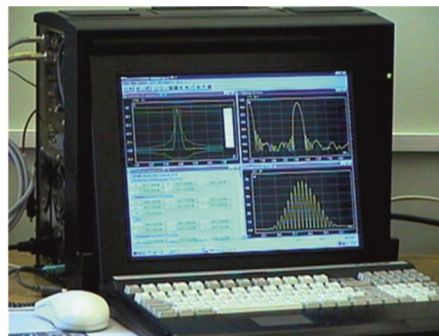


Рис. 16. Комплексы измерительные цифрового вещательного телевидения КИ-ТВМ и КИ-ЦВТ

Реализацию использования дистанционных телекоммуникационных систем можно пояснить на примере разработки телемедицинского комплекса МТМК.

Телемедицинские технологии, использующие видеоконференцсвязь, на протяжении последних лет доказали свою высокую эффективность.

Телемедицинский терминал видеоконференцсвязи должен обладать возможностью передачи в дуплексном режиме двух видеопотоков и непосредственного подключения медицинской диагностической аппаратуры по портам LAN, IEEE-1394, IEEE-802.11, USB, COM, Bluetooth, S-Video, RCA.

Комплексы должны быть оснащены необходимым и достаточным количеством медицинской диагностической аппаратуры и расходных материалов. При этом, набор медицинских диагностических методик и технических средств их реализации может различаться в зависимости от задач и условий применения.

Мобильный переносной телемедицинский комплекс на базе отечественных средств видеоконференцсвязи и отечественной медицинской диагностической аппаратуры может быть уже сейчас применён и для решения задач ведомственной медицины.

На рис. 17 приведена фотография опытного образца мобильного телемедицинского комплекса [29]. В качестве аппаратной вычислительной платформы выбран стандарт промышленных системных компьютерных плат mini-ITX, выпускаемых серийно многими известными компаниями на основе различных наборов системной логики (Intel, ATI, VIA, NVIDIA) и центральных процессоров (Intel, AMD, VIA). Выбор этого стандарта позволяет проводить модернизацию МТМК по мере выпуска более производительных процессоров с меньшим потреблением энергии без существенной переработки конструкторской документации.

Дисплей комплекса, смонтированный в крышке корпуса-чемодана, представляет собой 15-дюймовый сенсорный монитор разрешения XGA с вандалоустойчивым экраном из закалённого стекла. Это позволяет операто-

ру запускать программу видеоконференцсвязи, а также управлять ею простым касанием к пиктограмме на экране без пользования другими манипуляторами.



Рис. 17. Внешний вид телемедицинского комплекса

Важнейшей проблемой повышения эффективности современных разработок является анализ качества передаваемого и воспроизводимого мультимедийного контента.

До сих пор не существует единых методик оценки качества сервиса (QoS) и качества восприятия (QoE). В связи с этим проводятся разработки интеллектуальной методики оценки QoE с применением методов машинного обучения, позволяющей в автоматическом режиме оценивать качество восприятия на основе объективных параметров. Рассматриваются наборы анализируемых показателей, способы их получения и подходы к анализу.

Разрабатываемая методика обеспечит возможность оценивать QoE на основе различных наборов объективных данных, модульность методики позволяет адаптировать её в зависимости от специфики и возможностей исследуемых систем. При этом применяется вейвлет-анализ сетевого трафика, создаются математические модели особенностей контента.

В перспективе такой подход может стать универсальной методикой оценки качества восприятия мультимедийного контента.

Заключение

Рассмотренные в данной статье технические решения по внедрению современных телекоммуникационных технологий, включая отечественную систему РАВИС, должны способствовать реализации Указов Президента РФ № 203 от 9 мая 2017 г. «О Стратегии развития информационного общества в Российской Федерации на 2017 – 2030 годы» и № 208 от 13 мая 2017 г. «О стратегии экономической безопасности Российской Федерации на период до 2030 года».

Принимая во внимание стратегическую, идеологическую и научно-исследовательскую необходимость внедрения цифровой технологии РАВИС в РФ, считаем целесообразным принять следующие неотложные меры:

– Включить развертывание системы цифрового мультимедийного радиовещания РАВИС в реализацию программы «Цифровая экономика РФ», как одного из приоритетных направлений стратегического развития

страны в телекоммуникационной области до 2018 и на период до 2025 годов.

– В рамках реализации программы «Цифровая экономика РФ» реализовать бюджетное финансирование опытных зон системы мультимедийного вещания РАВИС в городах Ижевске, Казани, Калининграде и Краснодаре, обеспечив выделение частот в утвержденных ГКРЧ диапазонах для каждой опытной зоны и обеспечив выдачу разрешений на вещание.

Литература

1. «О Стратегии развития информационного общества в Российской Федерации на 2017 – 2030 годы». Указ № 203 Президента РФ Путина В.В. от 9 мая 2017 г.
2. Программа «Цифровая экономика Российской Федерации», утвержденная распоряжением Правительства Российской Федерации от 28 июля 2017 г. № 1632-р.
3. «О Стратегии экономической безопасности Российской Федерации на период до 2030 года». Указ № 208 Президента РФ Путина В.В. от 13 мая 2017 г.
4. Лейтес Л.С. Очерки истории отечественного телевидения. 2-е издание. М: Останкино, 2017, 223 с.
5. Лейтес Л.С. Телевидение России. Знаменательные факты, научно-техническая библиография, биографии известных ученых и специалистов техники телевидения. Справочник. М: Изд-во «Первый том», 2017, 580 с.
6. Report ITU-R BT.2380-0. Television colorimetry elements (07/2015).
7. Report ITU-R BT.2390-1. High dynamic range television for production and international programme exchange (10/2016).
8. Report ITU-R BS.2159-7. Multichannel sound technology in home and broadcasting applications (02/2015).
9. DIGITAG Digital Television Action Group. DTT Technology Information. DigiTAG - The Digital Terrestrial Television Action Group, 2015.
10. DIGITAG Digital Television Action Group. Guide to Digital Switchover. Focus Africa and Asia. DigiTAG - The Digital Terrestrial Television Action Group, 2013.
11. Framing Structure, Channel Coding and Modulation for Digital Television Terrestrial Broadcasting System (in Chinese). Chinese National Standard GB20600-2006.
12. Frame Structure, Channel Coding and Modulation for Digital Television/Terrestrial Multimedia Broadcasting-Advanced (in Chinese). Tsinghua University. National Engineering Lab. for DTV (Beijing) (DTVNEL).
13. Дворкович В.П., Дворкович А.В. Цифровые видеоинформационные системы (теория и практика) – М: Техносфера, 2012. – 1007 с.
14. Дворкович В.П., Дворкович А.В. Измерения в видеоинформационных системах (теория и практика) – М: Техносфера, 2015. – 873 с.
15. Дворкович В.П., Дворкович А.В. Оконные функции для гармонического анализа сигналов. Издание второе, переработанное и дополненное – М.: Техносфера, 2016. – 208 с.
16. Report ITU-R BT. 2295-1. Digital terrestrial broadcasting systems (07/2015).
17. Дворкович А.В., Дворкович В.П., Зубарев Ю.Б., Соколов А.Ю., Чернов Ю.А. Способ трансляции инфор-

мационного телевидения // Патент РФ № 2219676, приоритет от 08.11.2000.

18. Дворкович А.В., Дворкович В.П., Иртюга В.А. Способ мобильного узкополосного цифрового мультимедийного радиовещания // Патент РФ № 2441321, 27.01.2012, приоритет от 26.07.2010.

19. ГОСТ Р 54309-2011. Аудиовизуальная информационная система реального времени (РАВИС). Процессы формирования кадровой структуры, канального кодирования и модуляции для системы цифрового наземного узкополосного радиовещания в ОВЧ-диапазоне. Технические условия.

20. ГОСТ Р 55686-2013. Аудиовизуальная информационная система реального времени (РАВИС). Цифровой модулятор. Основные параметры и технические требования.

21. ГОСТ Р 55687-2013. Аудиовизуальная информационная система реального времени (РАВИС). Контрольный радиоприемник. Общие технические требования.

22. ГОСТ Р 55688-2013. Аудиовизуальная информационная система реального времени (РАВИС). Формирова-

тель контента. Структура и протоколы передачи данных.

23. ГОСТ Р 55689-2013. Аудиовизуальная информационная система реального времени (РАВИС). Нормы и методы метрологического обеспечения.

24. Report ITU-R BT.2049-7. Broadcasting of multimedia and data applications for mobile reception (02/2016).

25. Report ITU-R BS.2214-2. Planning parameters for terrestrial digital sound broadcasting systems in VHF bands (10/2016).

26. Свидетельство об утверждении типа средств измерений RU.C.35.002.A № 57904. Комплексы измерительные телевизионные КИ-ТВМ.

27. Свидетельство об утверждении типа средств измерений RU.C.35.002.A № 57905. Комплексы измерительные телевизионные КИ-ТВМ-Э.

28. Свидетельство об утверждении типа средств измерений RU.C.35.002.A № 57903. Комплексы измерительные цифрового вещательного телевидения КИ-ЦВТ.

29. Мобильный телемедицинский комплекс МТМК, Реклама НПФ «САД-КОМ», 2017.

НОВЫЕ КНИГИ

Брюханов Ю.А. Цифровые цепи и сигналы / Учебное пособие – М.: Изд-во «Горячая линия-Телеком», 3-изд. перераб. и доп., 2017 г. – 160 с.: ил.

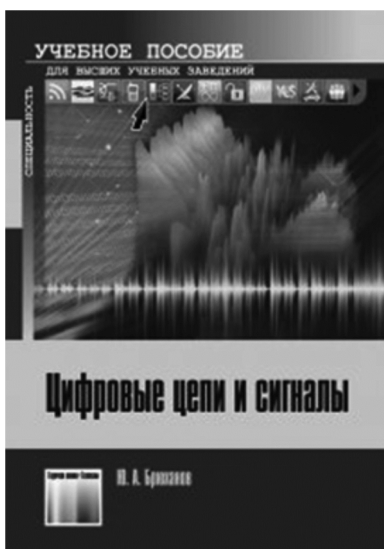
Приведен математический аппарат для анализа сигналов и цепей дискретного времени. Изложены спектральная теория периодических и непериодических цифровых сигналов, теория цепей дискретного времени. Подробно рассмотрены частотные свойства и временные характеристики базовых нерекурсивных и рекурсивных линейных цепей (цифровых фильтров) первого и второго порядков. Изложены методы изменения частоты дискретизации цифровых сигналов: децимация и интерполяция, а также способы их реализации. Уделено внимание теории случайных сигналов и процессов дискретного времени, протекающих в цифровых цепях. Рассмотрены эффекты квантования в цифровых сигналах и цепях.

Для студентов вузов, обучающихся по направлению подготовки «Радиофизика», «Радиотехника» и «Информационные технологии и системы связи».

Афанасьев А.А., Рыболовлев А.А., Рыжков А.П. / Учебное пособие для вузов – М.: Изд-во «Горячая линия-Телеком», 2017 г. – 356 с.: ил.

Изложены базовые теоретические и практические вопросы в области анализа, синтеза и экспериментального исследования систем цифровой обработки сигналов (ЦОС). Приведены основные методы расчета характеристик устройств ЦОС, показаны особенности их анализа в различных условиях функционирования, даны основные подходы к синтезу систем ЦОС, рассмотрены основные происходящие в них физические процессы. Изложенные теоретические сведения сопровождаются практическими примерами решения прикладных задач.

Для студентов вузов, обучающихся по укрупненной группе направлений подготовки 11.00.00 – «Электроника, радиотехника и системы связи»



УДК: 004.932.4

ВОССТАНОВЛЕНИЕ ФОНА В ВИДЕО С ИСПОЛЬЗОВАНИЕМ НЕПАРАМЕТРИЧЕСКОЙ МОДЕЛИ ДВИЖЕНИЯ И ПОКАДРОВОГО УТОЧНЕНИЯ

Боков А.А., аспирант лаборатории компьютерной графики и мультимедиа МГУ им. М. В. Ломоносова,
e-mail: abokov@graphics.cs.msu.ru;

Ватолин Д.С., к.ф.-м.н., с.н.с. лаборатории компьютерной графики и мультимедиа МГУ им. М. В. Ломоносова,
e-mail: dmitriy@graphics.cs.msu.ru.

VIDEO BACKGROUND RECONSTRUCTION BY NON-PARAMETRIC MOTION COMPLETION AND FRAME-BY-FRAME REFINEMENT

Bokov A.A., Vatolin D.S.

This paper introduces a novel video background reconstruction algorithm that combines the generality and the level of quality provided by the state-of-the-art approaches with much lower computational complexity, which is possible owing to per-frame processing instead of global spatio-temporal optimization. We achieve this result by using a general non-parametric method for motion completion and a frame-by-frame refinement step to alleviate the effect of error accumulation. The quality of the proposed approach is confirmed by the conducted objective evaluation and comparison with existing background reconstruction algorithms.

Key words: background reconstruction, motion completion, video processing.

Ключевые слова: восстановление фона, модель движения, покaдровая обработка, непараметрическая, вычислительная сложность.

Введение

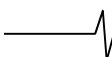
Восстановление фона является актуальной задачей интеллектуальной обработки видеопоследовательностей и имеет множество различных приложений. Оно применяется в киноиндустрии как в современных постановках для удаления нежелательных объектов из сцены, так и для восстановления архивных видеозаписей, содержащих дефекты пленки. Все большую популярность на сегодняшний день получает такой способ производства стереоскопических фильмов как 2D-3D конвертация, один из этапов которого заключается в восстановлении фона за объектами переднего плана, также известный как заполнение областей открытия. Более 60 % стереоскопических фильмов, вышедших в 2015-2016 годах в мировой прокат, были произведены именно путем конвертации из моноскопического формата. Также стоит упомянуть про такие приложения как маскирование ошибок видеокодека, вызванных потерей пакетов при передаче по сети [1], и заполнение неизвестных областей, возникающих при стабилизации видео с сохранением исходного разрешения [2].

Задача восстановления фона в видео, как правило, ставится следующим образом: на вход алгоритма помимо исходной видеопоследовательности подается бинарная маска неизвестной области Ω , где требуется восстановить фон на базе известной части входного видео V/Ω . Результатом работы алгоритма является восстановленный фон в области Ω , который оценивается по степени правдоподобности с точки зрения стороннего наблюдателя (обработанная видеопоследовательность не должна содержать видимых следов модификации). Данная задача достаточно широко исследована в

Представлен новый алгоритм восстановления фона в видео, который сопоставим с современными аналогами по уровню качества и ограничениям, накладываемым на входные данные, и при этом обладает существенно более низкой вычислительной сложностью за счет использования покaдровой обработки вместо решения глобальной задачи оптимизации по всему пространственно-временному объему. Данный результат обеспечивается использованием общей непараметрической модели движения фона и покaдрового уточнения результата, что предотвращает возможное накопление ошибки. Качество работы предложенного алгоритма подтверждается проведенным объективным сравнением с существующими аналогами.

литературе, однако до недавнего времени автоматические методы накладывали существенные ограничения на входную видеопоследовательность. Целый класс алгоритмов, известный как алгоритмы вычитания фона [3], опирается на предположение о статичности камеры для проведения восстановления. Многие предложенные алгоритмы используют простые параметрические модели для компенсации движения камеры, такие как аффинное [4], проективное [5] и кусочно-проективное [6] преобразования. Также многие алгоритмы [6, 7] делают предположение о том, что восстанавливаемые фрагменты фона должны быть видны хотя бы в одном кадре входного видео, что не всегда выполняется на практике. Современные подходы [8] во многом лишены вышеперечисленных недостатков, однако их применение существенно затрудняется высокой вычислительной сложностью и потреблением памяти, связанным с постановкой задачи восстановления фона как задачи оптимизации определенного функционала одновременно по всему пространственно-временному объему.

В данной работе представлен новый алгоритм восстановления фона в видео, основанный на использовании непараметрической модели движения и покaдровой оптимизации для уточнения результатов восстановле-



ния. Предложенный подход обладает значительно меньшей вычислительной сложностью, чем современные аналоги [4, 8] и позволяет обрабатывать видеопоследовательности с произвольным движением камеры и восстанавливать фрагменты фона, не видимые ни в одном из кадров входного видео. Качество работы предложенного алгоритма экспериментально подтверждено проведенным объективным сравнением с шестью аналогами, включая коммерческие инструменты для восстановления фона, с использованием тестовой выборки и методики оценки качества, описанной в [9]. Также проводится качественное сравнение предложенного алгоритма с современным аналогом [8] путем приведения результатов работы на репрезентативных кадрах.

Обзор области исследований

Одним из классических подходов к восстановлению фона в видео является использование трехмерных пространственно-временных блоков для оценки локального сходства восстановленного фона с известной частью входной видеопоследовательности V / Ω . В работе [10] задача формулируется как задача глобальной оптимизации по всему пространственно-временному объему, в ходе которой непосредственно минимизируется сумма расстояний от каждого трехмерного блока неизвестной области до наиболее схожего блока в известной части входного видео. Для нахождения наиболее схожего трехмерного блока в видео используется полный перебор и его простые оптимизации, что ограничивает область применимости метода до коротких последовательностей низкого разрешения (авторы приводят результаты на видео разрешения 360×288). В работе [4] был предложен способ значительного ускорения этапа поиска наиболее схожего трехмерного блока во входном видео путем использования алгоритма PatchMatch [12]. Авторами была модифицирована метрика расстояния между трехмерными блоками путем включения текстурных признаков, что позволило улучшить качество восстановления динамических текстур в видео, таких как волны, огонь, толпа людей. Однако, как было продемонстрировано в недавней работе [8], подходы, основанные на использовании трехмерных блоков, обладают существенными ограничениями при обработке видеопоследовательностей с произвольным движением камеры.

Некоторые алгоритмы накладывают явные ограничения на допустимый характер движения камеры и/или объектов во входном видео. В работе [13] авторы допускают лишь плоскопараллельное движение камеры во входной сцене. Авторы работы [14, 15] исходят из предположения, что входное видео содержит движущиеся объекты переднего плана и статичный фон. В [5] для компенсации движения камеры используется глобальное проективное преобразование, которое неспособно моделировать движение более чем одной плоскости в сцене при произвольном движении камеры. В работе [6] модель движения расширяется до кусочно-проективного преобразования (движение камеры компенсируется путем нахождения нескольких плоскостей в сцене и соответствующих им проективных преобразований). Авторы работы [6] также предлагают обобщение стандартного

метода бесшовного монтажа на случай восстановления фона в видео для компенсации возможных изменений освещенности. Однако данный алгоритм основан на попарном выравнивании всех входных кадров, что приводит к квадратичной зависимости времени работы алгоритма от количества кадров во входном видео. Авторы отмечают, что на обработку видео разрешения 960×720 у предложенного алгоритма уходит до 4 часов при работе на сервере с 64 логическими процессорами. В работе [16] используется схожий подход, но радиус поиска ограничивается соседними кадрами для снижения вычислительной сложности (авторы предлагают использовать скользящее окно, состоящее из 21 кадра). Несмотря на это, предложенный алгоритм сохраняет сравнительно высокую вычислительную сложность, которая затрудняет его применение на практике (авторы заявляют скорость работы в 1,5 минуты на кадр разрешения 1440×1056).

Наиболее схожими с предлагаемым в данной статье алгоритмом являются алгоритмы, основанные на непараметрических методах восстановления поля векторов движения в неизвестной области. В некоторых работах [11, 17, 18] восстановление поля векторов движения осуществляется путем копирования векторов из известной области с использованием трехмерных блоков для оценки локальной схожести поля векторов. Основным ограничением такого подхода является невозможность корректного восстановления поля векторов движения в неизвестной области простым копированием для некоторых типов движения камеры и изменений характеристик оптической системы объектива (например, масштабирование или наезд камеры). Авторы работы [19] применяют схожую технику с использованием трехмерных блоков, однако дополнительно выполняют итеративное уточнение поля векторов движения в неизвестной области для повышения качества. В работе [20] авторы сначала строят поле векторов движения в известной области и затем производят его покадровую интерполяцию в неизвестную область с использованием стандартного метода, основанного на диффузии. Далее непосредственно для восстановления фона авторы модифицируют метрику расстояния между двумерными блоками с учетом восстановленного поля векторов движения для поддержания временной стабильности. Поиск наиболее схожих двумерных блоков по всему входному видео, аналогично ранее рассмотренным алгоритмам, приводит к высокой вычислительной сложности (авторы сообщают, что предложенный алгоритм выполнил обработку видео разрешения 1280×720 продолжительностью 82 кадра за 3,3 часа). В другой работе [7] формулируется задача глобальной оптимизации для совместного нахождения восстановленного фона и соответствующего ему поля векторов движения. Авторы описывают итеративный алгоритм решения данной задачи, основанный на чередовании вычисления поля векторов движения во всей видеопоследовательности на базе текущего восстановленного фона и вычисления восстановленного фона на базе текущего поля векторов движения в неизвестной области Ω . Данный подход был далее развит в работе [8], где авторы демонстрируют высокое качество

восстановления фона на широком классе видеопоследовательностей с произвольным движением камеры и наличием фрагментов фона, не видимых ни в одном из кадров входного видео. Однако вследствие формулировки задачи как глобальной оптимизации по всему пространственно-временному объему, данный алгоритм сохраняет проблему подавляющего большинства существующих алгоритмов восстановления фона, которая заключается в чрезмерно высокой вычислительной сложности (предложенному авторами алгоритму требуется порядка 3 часов на обработку видео 854×480 продолжительностью 90 кадров).

Среди алгоритмов восстановления фона низкой вычислительной сложности стоит особо отметить работу [21], где предлагается специализированный алгоритм для восстановления фона в областях открытия, возникающих при генерации 3D-видео, основанный на локальной экстраполяции поля векторов движения. Однако данный алгоритм обладает рядом ограничений. Во-первых, использование локального копирования векторов движения в неизвестную область позволяет корректно восстанавливать движение фона лишь при плоско параллельном движении камеры и неспособно корректно обрабатывать сцены с существенным масштабированием или поворотом. Предложенный авторами алгоритм предполагает многократное копирование восстановленного фона вдоль восстановленных векторов движения с использованием бикубической интерполяции, что приводит к накоплению ошибки и размытию восстановленного фона, если он находится в рамках неизвестной области в течение длительного времени. Авторы работы [22] предлагают алгоритм восстановления фона, способный работать в режиме реального времени, однако данный алгоритм состоит в использовании глобального проективного преобразования для покадрового распространения результатов пространственного восстановления фона. То есть, предложенный авторами подход никак не использует фрагменты восстанавливаемого фона, видимые в других кадрах входной видеопоследовательности.

Таким образом, существующие алгоритмы восстановления фона либо накладывают ограничения на допустимые входные видеопоследовательности, либо обладают чрезмерно высокой вычислительной сложностью, что затрудняет использование соответствующих алгоритмов на практике. В то же время, предлагаемый в данной работе алгоритм не накладывает существенных ограничений на исходную видеопоследовательность, позволяя обрабатывать видео с произвольным движением камеры, благодаря использованию непараметрической модели движения; также он обладает более низкой вычислительной сложностью в сравнении с аналогами за счет применения покадровой оптимизации вместо преимущественно используемой в современных алгоритмах глобальной оптимизации по всему пространственно-временному объему.

Предлагаемый алгоритм восстановления фона

Как было отмечено ранее, отличительной особенностью предлагаемого алгоритма является покадровая

обработка входной видеопоследовательности. Алгоритм выполняет три прохода по исходной видеопоследовательности (в прямом направлении, в обратном, и снова в прямом), выполняя обработку очередного кадра с использованием результатов обработки предыдущего. В каждом из проходов неизвестная область в текущем кадре может быть восстановлена путем копирования фрагментов фона из любого из уже обработанных в данном проходе кадров (темпоральное восстановление). На отдельных кадрах также осуществляется восстановление фона путем копирования фрагментов из известной области того же самого кадра (пространственное восстановление). После каждого из проходов весь восстановленный фон фиксируется и используется как известный фон в следующем проходе. Обработка одного кадра в предложенном алгоритме включает в себя следующий набор шагов:

Восстановление поля векторов движения фона $O^{t \rightarrow t_0}$ внутри маски неизвестной области Ω^t от текущего кадра I^t к предыдущему обработанному кадру I^{t_0} (см. далее (1));

Построение поля векторов $V^t : \Omega^t \rightarrow R^2 \times N$, которое определяет отображение из неизвестной области на текущем кадре с номером t в известные области уже обработанных в данном проходе кадров. Для каждого пиксела в Ω^t данное поле векторов содержит либо относительный пространственный сдвиг и индекс исходного кадра, содержащего известный фон, либо особое значение V_{undef} , означающее, что в данном пикселе отображение в известный фон не определено. Данное поле векторов вычисляется с использованием результата предыдущего кадра V^{t_0} для построения начального приближения (см. далее (4));

Применение модифицированного алгоритма бесшовного монтажа для интеграции известных фрагментов фона, определенных полем векторов V^t , в текущий кадр с учетом возможных изменений условий освещенности сцены (см. далее (5));

Применение алгоритма пространственного восстановления фона на первом и последнем кадрах для обработки фрагментов неизвестной области, которые остаются скрытыми на протяжении всей входной видеопоследовательности.

Таким образом, предлагаемый алгоритм позволяет полностью использовать известный фон на всех кадрах входной видеопоследовательности, наряду с пространственным восстановлением на первом и последнем кадрах для фрагментов фона, не видимых ни в одном из кадров входного видео. Далее основные шаги алгоритма будут рассмотрены более подробно.

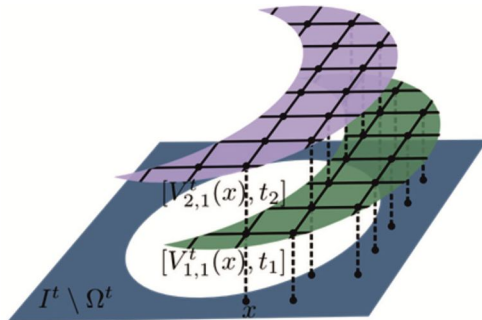
Восстановление поля векторов движения в неизвестной области

Существующие методы восстановления поля векторов движения либо используют стандартные алгоритмы построения оптического потока для известной области и отдельно формулируют задачу восстановления поля векторов движения в неизвестной области по известной [2, 17, 23], либо формулируют глобальную задачу опти-

мизации по всему видео для совместного нахождения цвета восстановленного фона и его движения [7, 8]. Отличительной особенностью предлагаемого в данной статье подхода является формулирование единой задачи оптимизации для совместного нахождения поля векторов движения в известной области и его гладкой интерполяции в неизвестной области с сохранением границ. Также, предложенный подход отличается от существующих алгоритмов построения оптического потока совмещением блочной оценки ошибки сопоставления и ограничений на гладкость поля векторов, характерных для вариационных алгоритмов построения оптического потока. А именно, поле векторов движения $O^{t \rightarrow t_0}$ от t -ого кадра до t_0 -ого вычисляется в узлах однородной сетки с шагом в s пикселей с использованием блоков $P(x, y)$ размера $s \times s$ для оценки ошибки сопоставления (в авторской реализации $s = 8$ пикселей) путем решения следующей оптимизационной задачи:

$$O^{t \rightarrow t_0} = \arg \min_O \sum_{(x, y)} \Phi \left(\sum_{\substack{p \in P(x, y) \\ p \notin \Omega^t \\ p+O(x, y) \in \Omega^0}} w_p (I^t(p) - I^{t_0}(p + O(x, y)) - M(x, y))^2 + \alpha \Phi(\|O(x+1, y) - O(x, y)\|^2 + \|O(x, y+1) - O(x, y)\|^2), \right. \\ \left. M(x, y) = \frac{\sum_p w_p (I^t(p) - I^{t_0}(p + O(x, y)))}{\sum_p w_p} \right) \quad (1)$$

Здесь: $\Phi(a^2) = \sqrt{a^2 + 0,001}$ – робастная функция потерь, позволяющая сохранить разрывы в поле векторов движения, $w_p = 1 / (1 + \|\nabla I^t(p)\|^2)$ – нормировочные веса, предотвращающие чрезмерное влияние областей с высокими значениями модуля градиента; $M(x, y)$ представляет собой нормировку на средние яркости блоков, что повышает устойчивость к изменениям освещения между соседними кадрами. Параметр α определяет степень гладкости получаемых полей векторов движения. В целях уменьшения вычислительной сложности при решении данной оптимизационной задачи



учитывается лишь яркостная компонента входной видеопоследовательности.

Аналогично существующим алгоритмам построения оптического потока [31], оптимизация данного функционала производится в рамках Гауссовой пирамиды с использованием разложения изображения в ряд Тейлора в предположении, что приращение вектора движения достаточно мало. Каждый следующий уровень пирамиды меньше предыдущего в 2 раза по каждой из осей; количество уровней пирамиды определяется разрешением исходного видео таким образом, чтобы нижний уровень пирамиды был последним, в который помещается хотя бы один блок $s \times s$ целиком. На каждом уровне пирамиды производится N_{outer} внешних итераций с разложением в ряд Тейлора до первого порядка включительно вокруг текущей аппроксимации вектора движения в каждом блоке. Из-за использования неквадратичной функции потерь, на каждую внешнюю итерацию также приходится N_{fp} итераций, в каждой из которых производится N_{SOR} итераций метода релаксации для решения соответствующей системы линейных уравнений (в авторской реализации $N_{outer} = 10$, $N_{fp} = 5$, $N_{SOR} = 25$). Финальное попиксельное поле векторов движения в неизвестной области получается путем билинейной интерполяции значений на разреженной сетке.

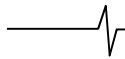
Построение поля векторов до исходных кадров

Первое приближение поля векторов $V^t: \Omega^t \rightarrow R^2 \times N$, определяющего пространственный сдвиг и номер исходного кадра с известным фоном для каждого пиксела неизвестной области Ω^t , может быть построено путем простого накопления векторов движения в неизвестной области, которые были вычислены на предыдущем шаге. А именно, первое приближение V^t строится путем прибавления поля векторов движения $O^{t \rightarrow t_0}$ к результату предыдущего кадра V^{t_0} :

$$V^t(x) = \begin{cases} [O^{t \rightarrow t_0}(x), t_0], & x_0 \notin \Omega^{t_0} \\ [O^{t \rightarrow t_0}(x) + V_1^{t_0}(x_0), V_2^{t_0}(x_0)], & V^{t_0}(x_0) \neq V_{undef} \\ V_{undef}, & \text{иначе.} \end{cases} \quad (2)$$

- – известный фон в текущем кадре
- – отображение в исходный кадр t_1
- – отображение в исходный кадр t_2
- ⋯ – ограничения на выравнивание фрагментов фона с разных исходных кадров
- ⋯ – ограничения на гладкость отображений в каждый исходный кадр

Рис. 1. Схема, иллюстрирующая расширенное определение поля векторов V^t , допускающее наличие нескольких отображений в различные исходные кадры в одном пикселе неизвестной области. На этапе уточнения поля векторов совместно оптимизируется выравнивание между фрагментами фона с разных кадров и гладкость отображений в каждый из них



Здесь: $x_0 = x + O^{t \rightarrow t_0}(x)$, $V_1^{t_0}(x_0)$ – первый элемент $V^{t_0}(x)$ (значение пространственного сдвига) и $V_2^{t_0}(x_0)$ – второй элемент (номер исходного кадра). Имея отображение в известный фон на другом кадре, определенное $V^t(x)$, можно восстановить фон в пикселе x простым копированием с билинейной интерполяцией. Однако экспериментально было установлено, что использование лишь восстановленных векторов движения в неизвестной области может приводить к накоплению ошибки. Для корректировки накопленной ошибки в данной работе предлагается алгоритм покадрового уточнения поля векторов V^t , основанный на оптимизации функционала, учитывающего как гладкость результирующего поля векторов, так и выравнивание между фрагментами фона, восстановленными с различных исходных кадров.

Для описания предложенного алгоритма уточнения поля векторов V^t введем ряд дополнительных обозначений. Во-первых, расширим определение V^t , чтобы оно допускало наличие нескольких отображений в известный фон на различных исходных кадрах в каждом пикселе $x \in \Omega^t$ (см. иллюстрацию на рис. 1). Точнее, будем предполагать, что пиксел x имеет $F^t(x)$ отображений в различные исходные кадры:

$$V^t(x) = \left\{ [V_{1,1}^t(x), V_{1,2}^t(x)], \dots, [V_{F^t(x),1}^t(x), V_{F^t(x),2}^t(x)] \right\}, \quad (3)$$

где $V_{i,1}^t(x), V_{i,2}^t(x)$ – это, соответственно, пространственный сдвиг и номер исходного кадра i -ого отображения. Следовательно, $V_{1,2}^t(x), \dots, V_{F^t(x),2}^t(x)$ образует множество различных номеров исходных кадров, в которые пиксел x имеет отображения. Также расширим область определения V^t до Ω_d^t – результата морфологического расширения Ω^t на d пикселов. С помощью данного обобщения мы можем ввести перекрытия между отображениями из неизвестной области в различные исходные кадры, что позволяет учитывать выравнивание между ними в процессе уточнения, причем ширина такого перекрытия становится параметром алгоритма d (в авторской реализации $d = 6$ пикселов). Таким образом, каждый пиксел в расширенной неизвестной области Ω_d^t содержит не более $d + 1$ отображений в различные исходные кадры. Поэтому описанную выше процедуру построения первого приближения V^t по результату предыдущего кадра V^{t_0} можно обобщить на расширенное определение V^t , допускающее перекрытия, с сохранением временной сложности пропорциональной количеству пикселов в Ω_d^t , вне зависимости от количества кадров во входной видеопоследовательности.

Сам алгоритм уточнения формулируется как задача оптимизации функционала, содержащего слагаемое $E_D(V^t)$, которое накладывает ограничение на выравнивание между фрагментами фона, восстановленными из различных исходных кадров, и слагаемое $E_S(V^t)$,

которое налагает штраф на разрывы в полях векторов, соответствующих отображениям в каждый из исходных кадров:

$$V^t = \arg \min_{V^t} (E_D(V^t) + \lambda E_S(V^t)), \quad (4)$$

$$E_D(V^t) = \sum_{x \in \Omega_d^t \setminus \Omega^t} \sum_{i=1}^{F^t(x)} \Phi((I^{V_{i,2}^t(x)}(x + V_{i,1}^t(x)) - I^t(x))^2) +$$

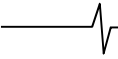
$$+ \sum_{x \in \Omega^t} \sum_{i=1}^{F^t(x)-1} \Phi((I^{V_{i+1,2}^t(x)}(x + V_{i+1,1}^t(x)) - I^{V_{i,2}^t(x)}(x + V_{i,1}^t(x)))^2),$$

$$E_S(V^t) = \sum_{(x_1, x_2) \in e(\Omega_d^t)} \sum_{\substack{1 \leq i_1 \leq F^t(x_1) \\ 1 \leq i_2 \leq F^t(x_2) \\ V_{i_1,2}^t(x_1) = V_{i_2,2}^t(x_2)}} \|V_{i_1,1}^t(x_1) - V_{i_2,1}^t(x_2)\|^2.$$

Здесь: $e(\Omega_d^t)$ – множество всевозможных пар соседних пикселов (x_1, x_2) , принадлежащих Ω_d^t . Первое слагаемое в $E_D(V^t)$ отвечает за выравнивание с известным фоном в текущем кадре, второе слагаемое отвечает за выравнивание фрагментов фона, восстановленных из различных исходных кадров, между собой. Оптимизация данного функционала проводится с помощью разложения в ряд Тейлора, аналогично оптимизации функционала (1). Для снижения вычислительной сложности также используется лишь яркостная компонента исходного видео. Однако здесь оптимизация происходит только на оригинальном разрешении, и разложение в ряд Тейлора применяется лишь один раз, так как использование восстановленного поля векторов движения фона уже предоставляет хорошее начальное приближение. Возникающая система линейных уравнений решается методом сопряженных градиентов с числом итераций $N_{CG} = 50$ в авторской реализации. Благодаря использованию результата предыдущего кадра для построения первого приближения поля векторов, даже такое малое число итераций позволяет получить решение приемлемого качества. В зависимости от структуры V^t количество неизвестных в СЛАУ может варьироваться, но его всегда можно ограничить сверху как $(d + 1) |\Omega_d^t|$.

Восстановление фона с помощью бесшовного монтажа

Наличие изменяющихся условий освещенности во входной видеопоследовательности или оставшиеся ошибки выравнивания между фрагментами фона, восстановленными с различных исходных кадров, могут приводить к образованию видимых швов в результате восстановления. Для снижения заметности швов в данной работе предлагается модификация метода бесшовного монтажа, известного как PoissonBlending [24]. Существующие обобщения этого метода для восстановления фона в видео [6] используют дополнительное слагаемое, накладывающее ограничение временной стабильности с учетом движения в восстановленной области. Отличительной особенностью предложенной модификации является использование адаптивных весов в слагаемом, отвечающим за поддержание временной



стабильности, что позволяет сохранять возможность равномерного изменения цвета всей восстанавливаемой области от кадра к кадру и, в то же время, усиливать ограничения временной стабильности в областях локально неконсистентного изменения цветов. А именно, предложенный алгоритм финального восстановления фона по вычисленному полю векторов V^t включает оптимизацию следующего функционала:

$$\begin{aligned}
 E(I^*) &= \\
 &= \sum_{p \in \Omega^t} \|\nabla I^*(p) - G^t(p)\|^2 + \sum_{p \in \Omega^t} w_p^{PB} \|I^*(p) - I^t(p)\|^2 + \\
 &+ \sum_{\substack{V^t(p) \neq V_{undef} \\ p \in \Omega^t}} (1 - w_p^{PB}) \|I^*(p) - I^{t_0}(p + O^{t \rightarrow t_0}(p))\|^2, \\
 w_p^{PB} &= \frac{1}{1 + \sigma_t \|\nabla I^{PB}(p) - G^t(p)\|^2}.
 \end{aligned} \tag{5}$$

Здесь, первое слагаемое штрафует отклонения поля градиентов результата от эталонного поля градиентов G^t , которое получается путем копирования градиентов из известных областей кадров в неизвестную область Ω^t , согласно полю векторов V^t . Второе слагаемое – мягкая версия жесткого граничного условия, требующего совпадение результата и исходного кадра на границе неизвестной области Ω^t . Третье слагаемое отвечает за поддержание временной стабильности относительно результата восстановления предыдущего обработанного кадра I^{t_0} с учетом восстановленного поля векторов движения $O^{t \rightarrow t_0}$. Компромисс между соблюдением граничных условий и поддержанием временной стабильности определяется весами w_p^{PB} , где I^{PB} – результат независимого применения метода PoissonBlending к текущему кадру (то есть, без учета третьего слагаемого и с $w_p^{PB} \equiv 1$). Таким образом, граничные условия сменяются на ограничения временной стабильности в областях с высокой ошибкой воспроизведения эталонного поля градиентов G^t , где изменение цвета, требуемое для сокрытия швов, существенно локально непостоянно. В ходе экспериментов было установлено, что такая модификация кардинально повышает устойчивость метода к локальным выбросам на отдельных кадрах, обеспечи-

вая высокую временную стабильность. Однако, в отличие от кадрового применения алгоритма PoissonBlending, резко возникающие локальные швы могут сохранять заметность в результате восстановления фона.

Для снижения вычислительной сложности вместо минимизации $E(I^*)$ напрямую проводится минимизация ее сепарабельной аппроксимации с использованием эвристики, предложенной в работе [25]. А именно, используя простое копирование согласно V^t для получения первого приближения восстановленного фона I_0 , чередуется независимое построчное решение с игнорированием всех вертикальных зависимостей и независимое решение по столбцам:

$$\begin{aligned}
 I'_k &= \arg \min_{I^*} \left(\epsilon_k \sum_p \|I^*(p) - I_{k-1}(p)\|^2 + E_H(I^*) \right), \\
 I_k &= \arg \min_{I^*} \left(\epsilon_k \sum_p \|I^*(p) - I'_k(p)\|^2 + E_V(I^*) \right),
 \end{aligned} \tag{6}$$

$k = 1, \dots, K$

Здесь: E_H и E_V – модификации исходного функционала E , игнорирующие все вертикальные и горизонтальные зависимости между пикселями решения, соответственно. Вычислительная сложность минимизации таких модификаций значительно ниже, так как она состоит в решении СЛАУ с трехдиагональной матрицей, а не пятидиагональной, как в случае минимизации исходного функционала. $\epsilon_k = \epsilon * a^{k-1}$ – экспоненциально возрастающие с номером итерации веса сохранения решения предыдущей итерации (в авторской реализации $\epsilon = 0,0001$, $a = 8$, количество итераций $K = 5$).

Пространственное восстановление фона

Для восстановления фрагментов, которые нельзя восстановить копированием с какого-либо из кадров входной видеопоследовательности, то есть, фрагментов, которые не видны ни в одном из кадров входного видео, используется пространственное восстановление фона. После обработки всех кадров в прямом порядке, к последнему кадру применяется алгоритм, предложенный в работе [26], для обработки фрагментов фона, которые остались невосстановленными. Затем, при обработке кадров в обратном порядке, восстановленный на последнем кадре фон будет использоваться алгоритмом

Таблица 1. Значения показателя качества MSDSSIM на тестовой выборке (меньшие значения соответствуют лучшему результату). TeleaInpainting и PlanarStructure являются алгоритмами пространственного восстановления фона

Алгоритм	Среднее место	Тестовая последовательность						
		№ 1	№ 2	№ 3	№ 4	№ 5	№ 6	№ 7
Предложенный	1,9	0,267	0,222	0,091	0,038	0,062	0,070	0,039
BGR [21]	2,6	0,221	0,192	0,217	0,070	0,090	0,108	0,083
F_RigRemoval [29]	2,7	0,291	0,211	0,078	0,120	0,068	0,091	0,104
PFClean[30]	3,4	0,307	0,187	0,077	0,094	0,143	0,163	0,106
ComplexScenes [4]	5,3	0,307	0,252	0,116	0,162	0,195	0,355	0,237
TeleaInpainting [27]	5,7	0,333	0,623	0,614	0,206	0,141	0,133	0,367
PlanarStructure [28]	6,4	0,318	0,603	0,682	0,240	0,177	0,302	0,432

Таблица 2. Значения показателя качества $C_{DSSIMdt}^{MS}$ на тестовой выборке (меньшие значения соответствуют лучшему результату). Данный показатель оценивает временную стабильность; методы пространственного восстановления (TeleaInpainting и PlanarStructure предсказуемо существенно проигрывают по этому параметру)

Алгоритм	Среднее место	Тестовая последовательность						
		№ 1	№ 2	№ 3	№ 4	№ 5	№ 6	№ 7
Предложенный	1,1	0,015	0,007	0,004	0,005	0,007	0,005	0,004
BGR [21]	2,7	0,015	0,007	0,024	0,009	0,009	0,010	0,007
PFClean [30]	3,0	0,017	0,007	0,004	0,011	0,015	0,013	0,009
F_RigRemoval [29]	3,6	0,020	0,015	0,006	0,014	0,008	0,009	0,011
ComplexScenes [4]	4,7	0,018	0,011	0,008	0,018	0,017	0,023	0,021
TeleaInpainting [27]	5,9	0,023	0,065	0,096	0,027	0,019	0,016	0,049
PlanarStructure [28]	7,0	0,071	0,105	0,128	0,060	0,039	0,081	0,091

Таблица 3. Время работы различных этапов предложенного алгоритма, описанных в соответствующих разделах статьи, и современного аналога на последовательности «satel» (90 кадров). Стоит отметить, что предложенный алгоритм работал в оригинальном разрешении 1920×1080, а время работы аналога указано авторами в статье [8] для уменьшенного разрешения 854×480

Алгоритм	Предложенный					Dynamic video [8]
	Формула (1)	Формула (4)	Формула (6)	Остальное	Целиком	
Время работы (сек.)	83,3	80,7	17,5	9,9	191,5	10800

как известная область, тем самым распространяя результат пространственного восстановления во времени. Аналогично, после обратного прохода алгоритма пространственное восстановление применяется к первому кадру видео. В результате, в ходе повторной обработки в прямом порядке восстановленным оказывается весь фон, находящийся в рамках кадра хотя бы на одном из крайних кадров входного видео. Для фрагментов фона, которые все еще остались не восстановленными после выполнения описанного алгоритма, применяется простой локальный метод пространственного восстановления [27].

Экспериментальная оценка

Для экспериментальной оценки предложенного алгоритма использовался набор из 7 тестовых видеопоследовательностей с известными эталонными результатами и два показателя качества из предложенных в работе [9]. А именно, было проведено сравнение предложенного алгоритма с аналогами с использованием показателя качества MSDSSIM, который оценивает схожесть восстановленного фона с эталоном напрямую, тем самым показывая достоверность результатов восстановления. Для оценки временной стабильности был выбран показатель $C_{DSSIMdt}^{MS}$, как имеющий наиболее высокую корреляцию с результатами экспертной оценки среди всех показателей, проанализированных в [9].

Во всех проводимых экспериментах значения параметров предложенного алгоритма были фиксированы: параметр гладкости оптического потока $\alpha = 2,5$, параметр гладкости в шаге уточнении поля векторов $\lambda = 200,0$, параметр временной стабильности в бесшовном монтаже $\sigma_t = 0,002$. По результатам проведенного сравнения предложенный алгоритм превзошел

аналоги как по достоверности восстановленного фона (см. табл. 1), так и по временной стабильности (см. табл. 2). Дополнительно сравним результаты работы предложенного подхода и алгоритма, предложенного в работе [8], на нескольких примерах (см. рис. 2); также проводится сравнение по скорости работы (см. табл. 3). Тестирование предложенного алгоритма проводилось на системе с Intel i5-6600K 3.5 GHz.

Как видно из приведенных результатов количественной и качественной оценки предложенного алгоритма, он не уступает современным аналогам по качеству восстановления фона и существенно их превосходит по скорости работы.

Заключение

В статье предложен новый алгоритм восстановления фона в видео, основанный на непараметрической модели движения фона и покадровом уточнении для предотвращения накопления ошибки, который не накладывает ограничений на допустимое движение камеры или объектов в сцене. Предложенный подход обладает существенно более низким временем работы, чем сопоставимые по качеству итогового результата аналоги, за счет вычисления оптического потока только в узлах разреженной сетки (1), использования приближенного итеративного метода решения СЛАУ с малым числом итераций при уточнении поля векторов (4) и сепарабельной аппроксимации в задаче бесшовного монтажа (6). В то же время, предложенный подход превзошел аналоги в проведенном объективном сравнении с алгоритмами с публично доступной реализацией и коммерческими инструментами для восстановления фона. Также, по качеству работы и широте класса допустимых входных видеопоследовательностей он сопоставим с современным подходом [8].



а) Сравнение на тестовой последовательности «camel»



б) Сравнение на тестовой последовательности «dance-twirl»



в) Сравнение на тестовой последовательности «rhino»

Рис. 2. Сравнение результата предложенного алгоритма и алгоритма, описанного в работе [8]. Авторы [8] дополнительно вручную модифицировали входные маски для удаления тени объектов

Литература

1. J. Liu, G. Zhai, X. Yang, B. Yang, and L. Chen, «Spatial error concealment with an adaptive linear predictor», IEEE Transactions on Circuits and Systems for Video Technology, pp. 353–366, 2015.
2. Y. Matsushita, E. Ofek, W. Ge, X. Tang, and H. Y. Shum, «Full-frame video stabilization with motion

inpainting», IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 1150-1163, 2006.

3. A. Sobral and A. Vacavant, «A comprehensive review of background subtraction algorithms evaluated with synthetic and real videos», Computer Vision and Image Understanding, pp. 4-21, 2014.

4. A. Newson, A. Almansa, M. Fradet, Y. Gousseau, and P. Pérez, «Video inpainting of complex scenes»,

SIAM Journal on Imaging Sciences, pp. 1993–2019, 2014.

5. Y. Shen, F. Lu, X. Cao, and H. Foroosh, «Video completion for perspective camera under constrained motion», IEEE International Conference on Pattern Recognition, pp. 63-66, 2006.

6. M. Granados, K. Kim, J. Tompkin, J. Kautz, and C. Theobalt, «Background inpainting for videos with dynamic objects and a free-moving camera», European Conference on Computer Vision (ECCV), pp. 682-695, 2012.

7. M. Roxas, T. Shiratori, and K. Ikeuchi, «Video completion via spatio-temporally consistent motion inpainting», IPSJ Transactions on Computer Vision and Applications, pp. 98-102, 2014.

8. J.-B. Huang, S. B. Kang, N. Ahuja, and J. Kopf, «Temporally coherent completion of dynamic video», ACM Transactions on Graphics (TOG), 2016.

9. Боков А.А., Ватолин Д.С. Методика объективной оценки качества восстановления фона в видео // Цифровая обработка сигналов. 2016. №3. С. 26-33.

10. Y. Wexler, E. Shechtman, and M. Irani, «Space-time video completion», IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), 2004.

11. T. Shiratori, Y. Matsushita, X. Tang, and S. B. Kang, «Video completion by motion field transfer», IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), pp. 411–418, 2006.

12. C. Barnes, E. Shechtman, A. Finkelstein, and D. Goldman, «Patchmatch: A randomized correspondence algorithm for structural image editing», ACM Transactions on Graphics (TOG), 2009.

13. K. Patwardhan and G. Sapiro, «Video inpainting under constrained camera motion», IEEE Transactions on Image Processing, pp. 545–553, 2007.

14. S.-C.-S. Cheung, J. Zhao, and M. V. Venkatesh, «Efficient object-based video inpainting», IEEE International Conference on Image Processing (ICIP), pp. 705-708, 2006.

15. M. Ebdelli, C. Guillemot, and O. Le Meur, «Exemplar-based video inpainting with motion-compensated neighbor embedding», IEEE International Conference on Image Processing (ICIP), pp. 1737-1740, 2012.

16. M. Ebdelli, O. Le Meur, and C. Guillemot, «Video inpainting with short-term windows: application to object removal and error concealment», IEEE Transactions on Image Processing, pp. 3034–3047, 2015.

17. M. Liu, S. Chen, J. Liu, and X. Tang, «Video completion via motion guided spatial-temporal global optimization», ACM International Conference on Multimedia, pp. 537-540, 2009.

18. N.-C. Tang, C.-T. Hsu, C.-W. Su, T.-K. Shih, and H.-Y.-M. Liao, «Video inpainting on digitized vintage films via maintaining spatiotemporal continuity», IEEE Transactions on Multimedia, pp. 602-614, 2011.

19. Z. Xu, Q. Zhang, Z. Cao, and C. Xiao, «Video Background Completion Using Motion-Guided Pixel Assignment Optimization», IEEE Transactions on Circuits and Systems for Video Technology, pp. 1393-1406, 2016.

20. M. Strobel, J. Diebold, and D. Cremers, «Flow and color inpainting for video completion», German Conference on Pattern Recognition (GCPR), pp. 293-304, 2014.

21. Зачесов А.А., Ерофеев М.В., Ватолин Д.С. Использование карт глубины при восстановлении фона в видеопоследовательностях // Новые информационные технологии в автоматизированных системах: материалы научно-практического семинара. 2015.

22. J. Herling and W. Broll, «High-quality real-time video inpainting with PixMix», IEEE Transactions on Visualization and Computer Graphics, pp. 866–879, 2014.

23. S. Liu, L. Yuan, P. Tan, and J. Sun, «Steadyflow: Spatially smooth optical flow for video stabilization», IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 4209-4216, 2014.

24. P. Pérez, M. Gangnet, and A. Blake, «Poisson image editing», ACM Transactions on Graphics (TOG), pp. 313-318, 2003.

25. D. Min, S. Choi, J. Lu, B. Ham, K. Sohn, and M.N. Do, «Fast global image smoothing based on weighted least squares», IEEE Transactions on Image Processing, pp. 5638-5653, 2014.

26. K. He and J. Sun, «Statistics of patch offsets for image completion», European Conference on Computer Vision (ECCV), pp. 16-29, 2012.

27. A. Telea, «An image inpainting technique based on the fast marching method», Journal of graphics tools, pp. 23–34, 2004.

28. J.-B. Huang, S. B. Kang, N. Ahuja, and J. Kopf, «Image completion using planar structure guidance», ACM Transactions on Graphics (TOG), 2014.

29. The Foundry Nuke. <https://www.thefoundry.co.uk/products/nuke/>.

30. Pixel Farm PFClean. <http://www.thepixelfarm.co.uk/pfclean/>.

31. P. Weinzaepfel, J. Revaud, Z. Harchaoui, and C. Schmid, «Deepflow: Large displacement optical flow with deep matching», IEEE International Conference on Computer Vision (ICCV), pp. 1385-1392, 2013.

РЕШЕНИЕ ПЕРЕОПРЕДЕЛЕННЫХ СИСТЕМ ЛИНЕЙНЫХ АЛГЕБРАИЧЕСКИХ УРАВНЕНИЙ В ЗАДАЧАХ ОБРАБОТКИ ИЗОБРАЖЕНИЙ

Новиков А.И., к.ф.-м.н., доцент кафедры «Высшая математика» Рязанского государственного радиотехнического университета, e-mail: novikovanatoly@yandex.ru.

THE SOLUTION OF OVERDETERMINED SYSTEMS OF LINEAR EQUATIONS IN PROBLEMS OF IMAGE PROCESSING

Novikov A.I.

We consider the problematic problems that arise when solving the overdetermined systems of linear algebraic equations in image processing problems. The simplest examples show the singularities of the solution of such systems by the method of least squares and the SVD method with the help of a singular expansion of the basic matrix of the system of equations.

Examples of solving overdetermined systems of equations in real image processing problems are given.

Practical recommendations are offered for finding pseudo-solutions of joint but uncertain systems of equations, as well as inconsistent in the classical sense of systems of equations.

Key words: image, homographic matrix, point cloud, singular matrix decomposition, SVD method.

Ключевые слова: изображение, матрица гомографии, облако точек, переопределенная система уравнений, нормальная система, псевдорешение, сингулярное разложение матрицы, метод SVD.

Введение

В самых различных областях приложений математики часто приходится решать системы линейных алгебраических уравнений (СЛАУ), число уравнений в которых существенно больше числа неизвестных. Такие системы уравнений называют переопределенными. Переопределенные СЛАУ часто возникают и при решении задач обработки изображений. Одна из них связана с нахождением матрицы

$$H = \begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & 1 \end{pmatrix} \quad (1)$$

(матрицы гомографии) преобразования точек $X' = \{(x'_i, y'_i)_{i=1}^n\}$ одного плоского изображения к плоскости другого $X = \{(x_i, y_i)_{i=1}^n\}$. В идеальном случае в однородных координатах должно выполняться равенство

$$\begin{pmatrix} \lambda x_i \\ \lambda y_i \\ \lambda \end{pmatrix} = \begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & 1 \end{pmatrix} \cdot \begin{pmatrix} x'_i \\ y'_i \\ 1 \end{pmatrix}. \quad (2)$$

В системе уравнений (1) неизвестными являются 8 коэффициентов h_{ij} матрицы гомографии. Записав в координатной форме уравнения (2) и исключив из них параметр λ , получим пару уравнений:

$$\begin{cases} x_i x'_i h_{31} + x_i y'_i h_{32} + x_i = x'_i h_{11} + y'_i h_{12} + h_{13} \\ y_i x'_i h_{31} + y_i y'_i h_{32} + y_i = x'_i h_{21} + y'_i h_{22} + h_{23} \end{cases} \quad (3)$$

Рассмотрены проблемные вопросы, возникающие при решении переопределенных систем линейных алгебраических уравнений в задачах обработки изображений. На простейших примерах показаны особенности решения таких систем методом наименьших квадратов и методом SVD с помощью сингулярного разложения основной матрицы системы уравнений.

Приведены примеры решения переопределенных систем уравнений в реальных задачах обработки изображений.

Предложены практические рекомендации по нахождению псевдорешений совместных, но неопределенных систем уравнений, а также, несовместных в классическом смысле систем уравнений.

После преобразования системы (3) к стандартной форме получим матричную форму записи СЛАУ

$$AG_H = B, \quad (4)$$

где $A = \begin{pmatrix} Z & O & W_1 \\ O & Z & W_2 \end{pmatrix}$ – основная матрица системы, в

составе которой блоки

$$Z = \begin{pmatrix} x'_1 & y'_1 & 1 \\ x'_2 & y'_2 & 1 \\ x'_3 & y'_3 & 1 \\ x'_4 & y'_4 & 1 \end{pmatrix}, \quad W_1 = \begin{pmatrix} -x_1 x'_1 & -x_1 y'_1 \\ -x_2 x'_2 & -x_2 y'_2 \\ -x_3 x'_3 & -x_3 y'_3 \\ -x_4 x'_4 & -x_4 y'_4 \end{pmatrix},$$

$$W_2 = \begin{pmatrix} -y_1 x'_1 & -y_1 y'_1 \\ -y_2 x'_2 & -y_2 y'_2 \\ -y_3 x'_3 & -y_3 y'_3 \\ -y_4 x'_4 & -y_4 y'_4 \end{pmatrix}, \quad O = (0)_{3 \times 3},$$

$G_H = (h_{11} \ h_{12} \ h_{13} \ h_{21} \ h_{22} \ h_{23} \ h_{31} \ h_{32})^T$ – вектор-столбец подлежащих определению элементов матрицы гомографии, $B = (x_1 \ x_2 \ x_3 \ x_4 \ y_1 \ y_2 \ y_3 \ y_4)^T$ – вектор-столбец координат ключевых точек первого изображения, записанных в указанном порядке.

Для нахождения оценок восьми коэффициентов необходимо как минимум 8 линейно независимых уравнений. Для получения 8 уравнений необходимо минимум 4 пары соответствующих (ключевых) точек [1]. Корректный выбор из множества пар ключевых точек четырех пар является достаточно сложной задачей, от успеха решения которой зависит не только качество совмещения изображений, но даже и сама возможность осуществления совмещения [2]. Названные причины и желание использовать максимальное число пар ключевых точек приводят к задаче поиска матрицы гомографии H по всему множеству найденных пар ключевых точек. В этом случае число n уравнений в СЛАУ (4) становится существенно больше числа неизвестных. В условиях ошибок измерений такая система уравнений – переопределенная СЛАУ – может оказаться несовместной в классическом смысле, и поэтому придется искать псевдорешение [3-6]. Аналогичная задача возникает при объединении (слиянии) двух облаков точек

$U = \{(x_i, y_i, z_i)_{i=1}^n\}$ и $V = \{(x'_i, y'_i, z'_i)_{i=1}^n\}$. Требуется найти такую матрицу

$$H = \begin{pmatrix} h_{11} & h_{12} & h_{13} & h_{14} \\ h_{21} & h_{22} & h_{23} & h_{24} \\ h_{31} & h_{32} & h_{33} & h_{34} \\ h_{41} & h_{41} & h_{41} & 1 \end{pmatrix}, \quad (5)$$

которая обеспечивала бы выполнение следующих матричных равенств в однородных координатах

$$\begin{pmatrix} \lambda x_i \\ \lambda y_i \\ \lambda z_i \\ \lambda \end{pmatrix} = \begin{pmatrix} h_{11} & h_{12} & h_{13} & h_{14} \\ h_{21} & h_{22} & h_{23} & h_{24} \\ h_{31} & h_{32} & h_{33} & h_{34} \\ h_{41} & h_{41} & h_{41} & 1 \end{pmatrix} \begin{pmatrix} x'_i \\ y'_i \\ z'_i \\ 1 \end{pmatrix}. \quad (6)$$

Из-за ошибок измерений матричное равенство (6) может не выполняться для некоторых точек из облаков точек. В этом случае по аналогии с рассмотренной выше двумерной задачей совмещения пары изображений по заданному множеству пар ключевых точек необходимо искать матрицу H , которая минимизировала бы расстояние между точками одного множества и преобразованного другого

$$H_{opt} = \arg \min_H \rho(X, HX').$$

Особенности решения переопределенных линейных систем

Проблемы, возникающие при решении переопределенных систем уравнений, удобно обсудить на простейших примерах. Рассмотрим СЛАУ

$$\begin{cases} x_1 + 2x_2 = 5, \\ 3x_1 + 4x_2 = 11. \end{cases}$$

Легко проверить, что она является совместной и переопределенной. Ее решением является вектор $X = (1; 2)^T$. Добавим к этой системе еще два уравнения

$$\begin{cases} x_1 + 2x_2 = 5, \\ 3x_1 + 4x_2 = 11, \\ 5x_1 + 6x_2 = 17, \\ 7x_1 + 8x_2 = 23. \end{cases} \quad (7)$$

Это пример переопределенной системы. Она также совместна и, по-прежнему, имеет единственное решение $X = (1; 2)^T$, так как третье и четвертое уравнения являются линейной комбинацией первых двух уравнений системы. Представим теперь, что в столбце свободных членов третий и четвертый коэффициенты претерпели незначительные изменения (см. систему (7))

$$\begin{cases} x_1 + 2x_2 = 5, \\ 3x_1 + 4x_2 = 11, \\ 5x_1 + 6x_2 = 16, \\ 7x_1 + 8x_2 = 24. \end{cases} \quad (8)$$

Такая ситуация является типичной при работе с реальными данными. Ошибки в регистрации результатов различных измерений неизбежны. Система уравнений (8) несовместна, так как не совпадают ранги ее основной и расширенной матриц (ранг основной матрицы равен 2, а расширенной 3). Вместе с тем из физических соображений, как правило, следует, что СЛАУ типа (8) должна иметь решение. В этом случае, как известно, ищут псевдорешение СЛАУ [3-6].

Задачи нахождения матриц гомографии при преобразовании одного множества точек к другому множеству точек с произвольным количеством точек в них как в двумерном случае (1), так и в трехмерном (5) формулируются одинаково

$$\|AG_H - B\|_{G_H}^2 \rightarrow \min. \quad (10)$$

В составе (10) $\|(\cdot)\|$ – евклидова норма. В двумерном случае структура матриц A, G_H, B приведена во введении. В трехмерном случае в задаче слияния облаков точек они будут иметь соответственно следующий вид:

A – блочная матрица размером $3n \times 15$:

$$A_{3n \times 15} = \begin{pmatrix} Z & O & O & W_1 \\ O & Z & O & W_2 \\ O & O & Z & W_3 \end{pmatrix}$$

с блоками в виде матриц $Z_{n \times 4} = (x_i, y_i, z_i, 1)_{i=1}^n$,

$$W_1 = (-x_i x'_i, -y_i x'_i, -z_i x'_i)_{i=1}^n, \quad W_2 = (-x_i y'_i, -y_i y'_i, -z_i y'_i)_{i=1}^n,$$

$$W_3 = (-x_i z'_i, -y_i z'_i, -z_i z'_i)_{i=1}^n;$$

O – нулевая матрица размером $n \times 4$,

G_H – вектор-столбец неизвестных размером 15×1 :

$$G_H = (h_{11} \ h_{12} \ h_{13} \ h_{14} \ h_{21} \ h_{22} \ h_{23} \ h_{24} \ h_{31} \ h_{32} \ h_{33} \ h_{34} \ h_{41} \ h_{42} \ h_{43})^T;$$

$$B = (x_1 \ x_2 \ \dots \ x_n \ y_1 \ y_2 \ \dots \ y_n \ z_1 \ z_2 \ \dots \ z_n)^T$$

размером $3n \times 15$.

В составе матрицы $H_{4 \times 4}$ (5) подматрица

$$R = \begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{pmatrix} - \text{отвечает за вращение облака}$$

точек, а подматрица $T = (h_{14} \ h_{24} \ h_{34})^T$ - за сдвиг облака точек.

Решение задачи (10) методом наименьших квадратов приводит к нормальной СЛАУ [7]

$$(A^T A)G_H = A^T B \quad (11)$$

с основной матрицей

$$A^T A = \begin{pmatrix} Z^T Z & O & Z^T W_1 \\ O & Z^T Z & Z^T W_2 \\ W_1^T Z & W_2^T Z & W_1^T W_1 + W_2^T W_2 \end{pmatrix}$$

при двумерной гомографии (2) и матрицей такой же структуры, но другого размера

$$A^T A = \begin{pmatrix} Z^T Z & O_{4 \times 4} & O & Z^T W_1 \\ O & Z^T Z & O & Z^T W_2 \\ O & O & Z^T Z & Z^T W_3 \\ W_1^T Z & W_2^T Z & W_3^T Z & \sum_{i=1}^3 W_i^T W_i \end{pmatrix},$$

при трехмерной гомографии (при преобразовании облаков точек).

Правые части нормальных уравнений при двумерной и трехмерной гомографии имеют соответственно следующий вид:

$$A^T B = \begin{pmatrix} Z^T B_1 & Z^T B_2 & \sum_{i=1}^2 W_i^T B_i \end{pmatrix}^T,$$

$$A^T B = \begin{pmatrix} Z^T B_1 & Z^T B_2 & Z^T B_3 & \sum_{i=1}^3 W_i^T B_i \end{pmatrix}^T.$$

В этих выражениях

$$B_1 = (x_1 \ x_2 \ \dots \ x_n)^T,$$

$$B_2 = (y_1 \ y_2 \ \dots \ y_n)^T, \quad - \text{вектор-столбцы размеров}$$

$$B_3 = (z_1 \ z_2 \ \dots \ z_n)^T$$

$n \times 1$.

В последние годы для решения переопределенных СЛАУ широко и часто без особых оснований используется, так называемый, метод SVD [4-6]. В его основе лежит сингулярное разложение основной матрицы СЛАУ в произведение левой и правой ортогональных матриц. Пусть

$$AX = B \quad (12)$$

СЛАУ, число m уравнений в которой больше числа n неизвестных. В методе SVD матрица A преобразуется в произведение матриц

$$A = USV^T. \quad (13)$$

В (13) матрицы U и V являются ортогональными, при этом матрица U называется левой сингулярной, а матрица V - правой сингулярной. Матрица S - диаго-

нальная матрица: $S = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_k, 0, 0, \dots, 0)$.

Числа $\sigma_i, i = \overline{1, k}$ называются сингулярными числами матрицы A . Отметим, что квадраты сингулярных чисел $\sigma_i, i = \overline{1, k}$ матрицы A совпадают с собственными значениями матрицы $A^T A$, а число k ненулевых значений сингулярных чисел равно рангу матрицы A .

Система уравнений (12) с учетом (13) принимает следующий вид $USV^T X = B$. Отсюда с учетом ортогональности матриц U и V получаем решение СЛАУ (12).

$$X = A^+ B = VS^{-1}U^T B. \quad (14)$$

Матрица $A^+ = VS^{-1}U^T$ в составе (13) называется псевдообратной, при этом

$$S^{-1} = \text{diag}\left(\frac{1}{\sigma_1}, \frac{1}{\sigma_2}, \dots, \frac{1}{\sigma_k}, 0, 0, \dots, 0\right).$$

Обращение к методу SVD при решении переопределенных СЛАУ не всегда оправдано. Если ранг основной матрицы в составе переопределенной и, возможно, несовместной СЛАУ (4) равен числу неизвестных ($\text{rg}A = n$), то решение СЛАУ (12) можно и целесообразно находить как решение нормальной СЛАУ (11). Решения нормальной СЛАУ (11), отвечающей СЛАУ (12) и решение СЛАУ (12) методом SVD в этом случае совпадают. Это, во-первых. Во-вторых, метод SVD требует большего числа вычислительных операций по сравнению с методом Гаусса, и, тем более, по сравнению с методом квадратного корня [7, 8], которым целесообразно решать нормальные СЛАУ, то есть СЛАУ с симметричной и положительно определенной основной матрицей [7].

Примеры решения переопределенных СЛАУ

Рассмотрим в качестве примера несовместную СЛАУ (8). Отвечающая ей нормальная СЛАУ (11) имеет следующий вид

$$\begin{pmatrix} 84 & 100 \\ 100 & 120 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 286 \\ 342 \end{pmatrix}. \quad (15)$$

Решением этой совместной системы и одновременно псевдорешением СЛАУ (8) будет вектор $X = (1,5 \ 1,6)^T$.

Методом SVD система уравнений (8) приводится к виду

$$\begin{pmatrix} 0,1525 & 0,8226 \\ 0,3499 & 0,4214 \\ 0,5474 & 0,0201 \\ 0,7448 & -0,3812 \end{pmatrix} \begin{pmatrix} 14,269 & 0 \\ 0 & 0,626 \end{pmatrix} \times \begin{pmatrix} 5 \\ 11 \\ 16 \\ 24 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 5 \\ 11 \\ 16 \\ 24 \end{pmatrix} \quad (16)$$

В левой части системы (16) первые три матрицы являются соответственно левой сингулярной матрицей U , диагональной матрицей сингулярных чисел S и правой

сингулярной матрицей V^T . Сингулярное разложение матрицы A в составе (16) иногда называют экономным сингулярным разложением. В полном сингулярном разложении матрицы U и S будут иметь размер 4×4 , а матрица V^T - 4×2 . Недостающие два столбца и две строки в составе матриц U и V^T можно найти с помощью алгоритма ортогонализации Грамма-Шмидта [9].

Решением СЛАУ (16) и, значит, псевдорешением СЛАУ (8) является вектор $X = (1,5 \ 1,599)^T$. Как и следовало ожидать, решение нормальной СЛАУ (15) и решение СЛАУ (8) методом SVD (решение СЛАУ (16)) практически совпали.

Рассмотрим еще один пример

$$\begin{cases} x_1 - 2x_2 + x_3 = 2, \\ x_1 + x_2 + x_3 = -1, \\ x_1 + 4x_2 + x_3 = -3. \end{cases} \quad (17)$$

Ранги основной и расширенной матриц данной СЛАУ не совпадают ($rgA = 2$, $rg(A|B) = 3$). Значит, в классическом смысле данная система уравнений несовместна. Найдем псевдорешение СЛАУ (17).

Нормальная СЛАУ, отвечающая этой системе уравнений, имеет вид

$$\begin{pmatrix} 3 & 3 & 3 \\ 3 & 21 & 3 \\ 3 & 3 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -2 \\ -17 \\ -2 \end{pmatrix}.$$

Ранги основной и расширенной матриц этой системы равны 2. Значит, эта СЛАУ совместна, но ее решением будут все точки, лежащие на некоторой прямой в пространстве R^3 . Искомым решением является вектор

$$X = \left(\frac{1}{6} - t \quad -\frac{5}{6} \quad t \right)^T.$$

Минимизация полученного решения по параметру t приводит нас к нормальному псевдорешению $X = (0,08(3) \ 0,8(3) \ 0,08(3))^T$ системы.

Отметим, что «псевдорешение»

$X = (-1,3001 \ -0,0000 \ 1,3001)^T$ СЛАУ (17), найденное методом SVD с помощью стандартной программы в MATLAB, существенно отличается от найденного нормального псевдорешения. При этом программа предупреждает о возможной ошибке (Warning: Matrix is close to singular or badly scaled. Results may be inaccurate.). Следует отметить, что псевдорешение системы (17), найденное методом SVD с помощью оригинальной программы (авторы Колчаев Д.А., Новиков А.И.) практически совпадает с нормальным псевдорешением: $X = (0,0835 \ 0,8340 \ 0,0835)^T$.

Практические выводы из этих двух примеров. Если а priori известно, что основная матрица A переопределенной СЛАУ (12) является матрицей полного ранга ($rgA = n$), то целесообразно решение (нормальное «псевдорешение») системы искать как решение нормальной СЛАУ (11). Если же известно, что ранг основной матрицы A переопределенной СЛАУ (12) может

быть меньше числа неизвестных ($rgA = k < n$), то псевдорешение системы (12) целесообразно искать методом SVD. Хотя, заметим, что и в этом случае можно искать нормальное псевдорешение системы, но с дополнительной минимизацией найденного $(n - k)$ - параметрического решения по этим параметрам.

Пример решения СЛАУ (17) различными методами и с помощью различных вариантов программ подтверждает верность еще одного утверждения: нельзя отождествлять алгоритм решения задачи и некоторую его программную реализацию.

Рассмотрим примеры нахождения нормального решения СЛАУ и решения (псевдорешения) этих же СЛАУ с помощью метода SVD в реальных задачах [2, 10]. В [2] приведены вычисленные в результате решения нормальной СЛАУ (11) матрицы гомографии в одном случае по 18 парам ключевых точек, а в другом – после удаления некорректной пары точек – по 17 парам точек. Во втором случае (по 17 парам точек) матрица гомографии, приведенная в [2], имела следующий вид

$$\begin{pmatrix} 0,941 & 0,052 & 21,93 \\ 0,047 & 1,018 & 17,86 \\ 8,32 \times 10^{-6} & 3,74 \times 10^{-5} & 1 \end{pmatrix}.$$

Матрицы гомографии, полученные в результате решения переопределенных систем уравнений (18 и 17 уравнений относительно 8 неизвестных) методом SVD, полностью совпали с решениями нормальных СЛАУ (11). Это ожидаемый результат, поскольку в обоих случаях основная матрица СЛАУ была матрицей полного ранга.

Рассмотренный частный пример на самом деле является подтверждением общего свойства алгоритмов построения проективных преобразований: основная матрица СЛАУ типа (4) является матрицей полного ранга: $rgA = \min\{m, 8\} = 8$. Поэтому решение нормальной СЛАУ (11) и решение СЛАУ (4) методом SVD в задачах построения проективных преобразований должны совпадать.

В общем случае нормальное псевдорешение СЛАУ (12) и псевдорешение, найденное с помощью сингулярного разложения основной матрицы системы, будут совпадать, если будет выполнено условие $rgA = \min\{m, n\} = n$ ($n \ll m$).

Преимущество получения нормального решения либо нормального псевдорешения в случае несовместной СЛАУ по сравнению с методом SVD заключается в меньшем объеме вычислительных процедур.

Положительный результат решения соответствующих СЛАУ методом SVD заключается в том, что сингулярное разложение матрицы A позволяет лучше понять особенности системы уравнений. Так, в рассмотренном примере (16) из работы [2] первые два сингулярных числа в матрице S оказались существенно больше остальных шести сингулярных чисел. Значит, экономное сингулярное разложение матрицы A с левой сингулярной матрицей U размером 8×2 , матрицей $S_{2 \times 2}$ и матрицей $V_{2 \times 2}^T$ уже позволило бы получить приближенное решение задачи.

Следующий пример построения нормального решения (псевдорешения) переопределенной СЛАУ связан с совмещением двух облаков точек. Первое облако точек является поверхностью гиперболического параболоида

$$z = \frac{x^2}{50} - \frac{y^2}{75}, \quad x = \overline{-10,20}; \quad y = \overline{-8,8}.$$

Оно изображено в нижней части рис. 1.

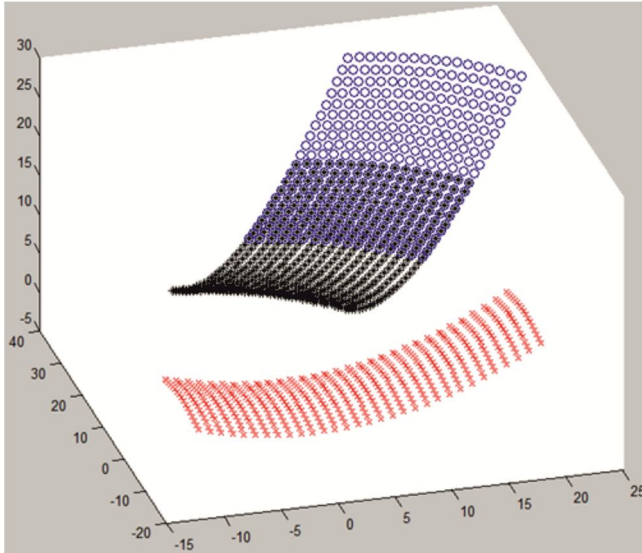


Рис. 1. Результат совмещения облаков точек ($\sigma = 0,01$)

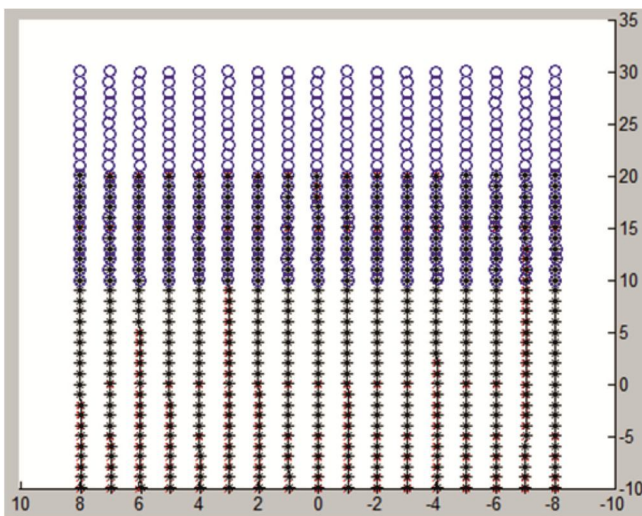


Рис. 2. Результат совмещения облаков точек в проекции точек на плоскость Oxy ($\sigma = 0,01$)

Второе облако точек получено по аналогичной формуле

$$z = \frac{x^2}{50} - \frac{y^2}{75}, \quad x = \overline{10,30}; \quad y = \overline{-8,8}.$$

Первое и второе облако точек имеют общую часть (область перекрытия) в количестве 187 точек: по координате x от 10 до 20, а по координате y – от -8 до 8. Она хорошо видна на рис. 2 и рис. 4 как горизонтальная полоса из совмещенных точек первого и второго облаков в промежутке от 10 до 20 по оси Oy . После формирования облаков точек второе облако подверглось преобразованиям сдвига и поворота. По координате z был выполнен сдвиг на 15 единиц. Затем оно было повернуто относительно каждой из трех осей Ox , Oy и Oz на

углы соответственно $6^\circ, 5^\circ, 60^\circ$. После этого к каждой координате «искаженного» первого облака точек последовательно прибавлялся гауссовский дискретный шум увеличивающейся интенсивности.

В верхней части рис. 1 представлено множество точек второго облака точек после преобразований сдвига и вращений относительно координатных осей с искажениями гауссовским шумом с СКО $\sigma = 0,01$ по каждой координате. Здесь же (в верхней части рисунка) приведен результат слияния облаков точек.

Система линейных алгебраических уравнений, вытекающая из матричного равенства (8), состоит из 187 уравнений относительно 15 неизвестных – элементов матрицы гомографии в трехмерном случае – и является переопределенной СЛАУ. Нормальное псевдорешение этой системы находилось как решение СЛАУ (11).

Матрица гомографии, обеспечивающая «слияние» облаков точек (преобразование первого облака точек ко второму) для $\sigma = 0,01$ имеет следующий вид

$$\begin{pmatrix} 0,997 & 6,31 \cdot 10^{-5} & 0,005 & 0,004 \\ 0,001 & 0,999 & -0,001 & -0,015 \\ -0,005 & -1,41 \cdot 10^{-5} & 1,009 & 15,011 \\ 2,35 \cdot 10^{-4} & 1,14 \cdot 10^{-5} & 5,18 \cdot 10^{-4} & 1 \end{pmatrix}$$

Результат слияния облаков точек с помощью этой матрицы гомографии приведен в верхней части рис. 1. Совмещение почти идеальное. Это особенно хорошо видно в проекции точек из совмещенных облаков на плоскость Oxy (рис. 2).

Уклонение $\rho(HX, X')$ образов точек HX первого облака от отвечающих им точек второго облака в области перекрытия $x = \overline{10,20}; \quad y = \overline{-8,8}$ равно $\rho(HX, X') = 0,008$.

Увеличение интенсивности шума приводит к существенным искажениям преобразованного облака точек. Так при $\sigma = 0,05$ результат совмещения облаков точек является вполне удовлетворительным в области перекрытия изображений (рис. 3), и далек от удовлетворительного на промежутке $x \in [-10, 10], \quad y \in [-8, 8]$.

Это хорошо видно в проекции совмещенных облаков точек на плоскость Oxy (рис.4)

Эти искажения обусловлены изменениями элементов матрицы гомографии

$$\begin{pmatrix} 0,890 & 0,002 & 0,100 & 0,578 \\ -0,002 & 0,957 & 0,003 & 0,014 \\ -0,087 & -1,41 \cdot 10^{-5} & 1,009 & 15,125 \\ -0,005 & 1,36 \cdot 10^{-5} & 0,007 & 1 \end{pmatrix}$$

Проблема точности слияния облаков точек не связана прямо с обсуждаемыми в данной статье методами решения переопределенных СЛАУ, но, естественно, заслуживает отдельного обсуждения. Приведем основные публикации по этой проблеме. Широко используемым в практике совмещения трехмерных моделей (облаков точек) является итеративный алгоритм ближай-

ших точек ICP [11] с последовательным уточнением соответствия между двумя облаками точек. Алгоритм состоит из двух основных этапов: поиск соответствующих точек в двух облаках на первом этапе, определение параметров преобразования (растяжение, сдвиг и вращение), применение этого преобразования и оценивание функции ошибки – на втором этапе. В качестве функции ошибки в методе point-to-point используется расстояние между соответствующими точками в паре облаков, а во втором методе point-to-plane – расстояние между точкой одного облака и касательной плоскостью к другому облаку.

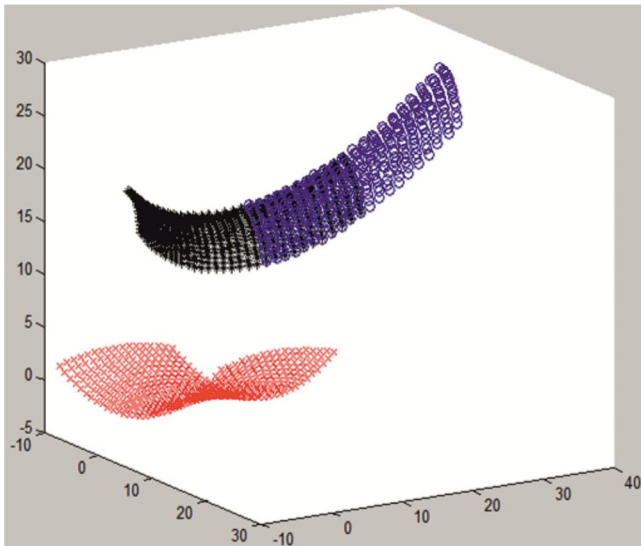


Рис. 3. Результат совмещения облаков точек ($\sigma = 0,05$)

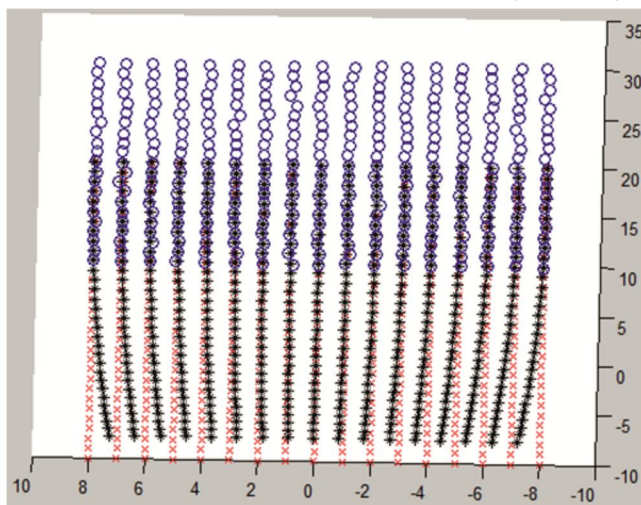


Рис. 4. Результат совмещения облаков точек в проекции точек на плоскость Oxy ($\sigma = 0,05$)

Алгоритм имеет многочисленные модификации, например [12, 13]. В работе [14] предложено замкнутое решение задачи слияния двух облаков точек с функцией ошибки, вычисляемой по методу point-to-plane.

Заключение

Рассмотрены проблемы, возникающие при решении переопределенных и возможно несовместных в класси-

ческом смысле систем линейных алгебраических уравнений и методы решения таких систем в задачах обработки изображений. На простейших примерах проведено сопоставление решения таких систем методом наименьших квадратов с выходом на нормальную систему уравнений и нормальное псевдорешение, с одной стороны, и методом SVD сингулярного разложения основной матрицы системы, с другой.

Приведены примеры переопределенных СЛАУ, возникающих в задачах обработки изображений и их решение названными альтернативными методами.

Предложены практические рекомендации по использованию методов решения линейных систем уравнений в задачах обработки изображений.

Литература

1. A.I. Novikov, V.A. Sablina, M.B. Nikiforov and A.A. Loginov. The Contour Analysis and Image-Superimposition Problem in Computer Vision Systems // Pattern Recognition and Image Analysis, 2015, Vol.25, No 1, pp. 73-80. - DOI: 10.1134/S1054661815020194.
2. Ефимов А.И. Алгоритм поэтапного уточнения проективного преобразования для совмещения изображений / А.И. Ефимов, А.И. Новиков // Компьютерная оптика. – 2016. – Т. 40, № 1. – С. 258-266. – DOI: 18287/2412-6179-2016-40-2-258-266.
3. Воеводин В.В., Кузнецов Ю.А. Матрицы и вычисления. М.: Наука, 1984. – 320 с.
4. Андрушевский Н.М. Анализ устойчивости решений систем линейных алгебраических уравнений. М.: Издательский отдел факультета ВМиК МГУ им. М.В. Ломоносова; МАКС Пресс, 2008. – 71 с.
5. Деммель Дж. Вычислительная линейная алгебра. Теория и практика. М.: Мир, 1999. – 430 с.
6. Голуб Дж., Ван Лоун Ч. Матричные вычисления. М.: Мир, 1999. – 548 с.
7. Демидович Б.П., Марон И.А. Основы вычислительной математики. М.: Наука, 1966. – 664 с.
8. Икрамов Х.Д. Численные методы для симметричных линейных систем. М.: Наука, 1988. – 160 с.
9. Ильин В.А., Позняк Э.Г. Линейная алгебра. М.: Наука, 1984. – 294 с.
10. Холопов И.С. Алгоритм коррекции проективных искажений при маловысотной съемке / И.С. Холопов // Компьютерная оптика. -2017. –Т. 41, № 2. – с. 284-290. – DOI: 10.18287/0134-2452-2017-41-2-284-290.
11. Besl P.J., McKey, Neil D. A Method for Registration of 3-D Shapes // IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 14, № 26 1992. pp. 239-256.
12. Du S. Affine iterative closest point algorithm for point set registration/ S. Du, N. Zheng, S. Ying, J. Liu // Pattern Recognition Letters/ - 2010 – 31/ -P. 791-799.
13. A. Makovetskii, S. Voronin, V. Kober, D. Tihonkih. Affine registration of point-to-plane approach//Сб. трудов III межд. Конф. и молодежной школы ITNT – 2017, Самара: Новая техника. 2017. с. 684-688.

УЛУЧШЕНИЕ ПРОЦЕССА ТЕМАТИЧЕСКОЙ ОБРАБОТКИ ГИПЕРСПЕКТРАЛЬНОЙ ИНФОРМАЦИИ

Третьяков В.А., главный специалист ФГУП «Центральный научно-исследовательский институт машиностроения» (ФГУП ЦНИИмаш), e-mail: tretyakovva@tsniimash.ru;

Кротков А.Ю., главный специалист ФГУП «ЦНИИмаш», e-mail: krotkovay@tsniimash.ru;

Кривошеин В.В., ведущий инженер ФГУП «ЦНИИмаш», e-mail: krivosheinvv@tsniimash.ru;

Данилов Р.Ю., к.б.н., старший научный сотрудник ФГБНУ «Всероссийский научно-исследовательский институт биологической защиты растений», e-mail: daniloff.roman@yandex.ru.

IMPROVING THE THEMATIC PROCESSING OF HYPERSPECTRAL INFORMATION

Tretiakov V.A., Krotkov A.Y., Krivoshein V.V., Ph.D. Biology Danilov R.Y.

Thematic processing of hyperspectral information is based on methods of statistical pattern recognition. In this paper, improving the process of the recognition of two classes of ground objects to reduce number and time of calculations is presented. For this purpose we consider the lower recognition error probability estimate limit by spectral features (reflectance). The main assumptions are: classification of the test sample, the equality of covariance matrices of two classes, the independence of spectral features.

Methods for determining the minimum required values of the recognition parameters of the selected object classes are developed under the given assumptions. Method based on one sample t-test is developed to determine the values of wavelengths corresponding to the spectral features, under the made assumptions.

Ground experimental hyperspectral measurements, carried out on the test site of Krasnodar, were done in order to verify the developed methods.

The minimum required volume of the sample, based on two sample Fisher's test is calculated. The necessary number of the spectral informative features is defined using given recognition error probability of two classes of ground objects.

Key words: hyperspectral measurements, reflectance, spectral features, statistical pattern recognition, error probability, training data.

Ключевые слова: гиперспектральные измерения, коэффициент спектральной яркости, спектральные признаки, статистическое распознавание, вероятность ошибки, обучающая выборка.

Введение

В настоящее время в нашей стране гиперспектральное (ГС) дистанционное зондирование Земли (ДЗЗ) является активно развивающимся направлением. Интерпретация данных ГС ДЗЗ и решение тематических природоресурсных задач требует разработки сложных алгоритмов, в основе которых лежат методы статистического распознавания объектов. Существует множество отечественных и зарубежных публикаций, в которых описываются подобные методы, а также результаты применения их в тематической обработке ГС данных. Однако, на сегодняшний день недостаточно проработан вопрос обоснования объема обучающей выборки в случае рассмотрения методов контролируемой классификации. В работе [1], посвященной классификации сельскохозяйственных культур с помощью гиперспектральных данных НУМАР, выбирается два набора данных, каждый из которых содержит 6 классов. Объем обучающей выборки из каждого класса выбирается равным 15 % от общего количества пикселей. Первый из двух наборов данных используется для обучения алгоритма классифи-

Предложен подход к улучшению процесса статистического распознавания двух классов наземных объектов с оценкой нижнего предела вероятности ошибки. Проведены расчеты с использованием экспериментально отработанных методик получения и обработки коэффициента спектральной яркости растительности на основе опыта многолетних наземных спектральных измерений, проведенных на Краснодарском тестовом участке для отработки разработанного критерия Фишера и количество спектральных информативных признаков для распознавания объектов с заданной точностью, и оценкой нижней границы вероятности ошибки. С учетом сделанных допущений на основе одно выборочного t-критерия разработан метод определения значений длин волн, соответствующих спектральным признакам.

кации методом максимума правдоподобия. Второй набор данных необходим для оценки точности классификации данных. Рассматривался вопрос определения необходимого количества спектральных признаков и был предложен метод на основе вычисления максимального детерминанта ковариационной матрицы классов. Недостаток метода заключается в необходимости перебора каждого спектрального признака, пока добавление нового признака не перестанет увеличивать детерминант ковариационной матрицы.

В работе [2] классифицировались такие объекты, как реки, лес, поля с овощами, крыши домов, земли с за-

стройкой, на мультиспектральном изображении (6 каналов), полученном с оптического датчика ТМ космического аппарата Landsat различными методами классификации с использованием 12 наборов обучающих выборок объемами 20, 40, 60, 80, 100, 120, 140, 160, 180, 200, 220, 240.

Еще в одной работе [3] исследовалась точность классификации шести классов объектов на изображении высокого разрешения с оптического сенсора космического аппарата Quickbird методом опорных векторов, методом расстояния Махаланобиса, методом максимума правдоподобия и т.д. при различных объемах обучающей выборки 100, 200, 300, 400 для каждого класса. Наивысшая точность 80 % достигается при использовании метода опорных векторов при размере обучающей выборки 200. Метод максимума правдоподобия обеспечил точность распознавания 78,33 % при том же объеме обучающей выборки.

Во всех работах исследовалось влияние объема обучающих выборок на точность классификации хорошо различимых объектов (водные объекты, почвенные объекты, растительность, здания, дороги и т.д.). В нашей работе рассматривается процесс тематической классификации двух классов объектов с тонкими спектральными различиями, описанными в книге [4], и для его улучшения разрабатываются методы определения минимально необходимых параметров распознавания: объем выборки и количество информативных признаков на основе гиперспектральных измерений.

Постановка задачи исследования

Пусть на основе анализа процесса А в составе процесса Б, где А – процесс тематической классификации двух классов наземных объектов, Б – тематическая обработка ГС изображения, определены:

а) совокупность X параметров элемента А, где X – гиперспектральные данные;

б) показатель эффективности P_e элемента А, где P_e – вероятность ошибки классификации;

в) параметры W , оказывающие наибольшее влияние на выбранный показатель, где W – совокупность таких параметров, как объем обучающей выборки v , необходимое количество спектральных каналов n , значения длин волн w ;

г) система основных ограничений и допущений D , принятая при проведении исследований элемента А, подробно описана ниже.

В ходе исследований требуется разработать математическую модель, которая бы позволила:

1) установить связь

$$P_e = P_e(X, W, D)$$

показателя P_e с параметрами $X, (n, v, w) \in W$, при заданных значениях D ;

2) определить n при заданных значениях P_e и D ;

3) определить v при заданных значениях D ;

4) определить w при заданных значениях P_e и D .

Методы решения задачи исследования

В основе тематической обработки гиперспектральной информации лежат методы статистического распо-

знавания объектов. Признаками распознавания являются коэффициенты спектральной яркости (КСЯ) наземных объектов, по которым строятся нормальные распределения рассматриваемых классов. Отдельный класс представляет собой набор КСЯ, каждый из которых свернут в одну точку [5]. Предположим, что два выбранных класса наземных объектов, являющиеся множеством КСЯ, подчиняются нормальному закону распределения. Для этого случая приведем систему условий D оценки нижней границы вероятности ошибки статистического распознавания двух классов объектов со своими многомерными нормальными распределениями яркости по спектральным признакам:

1) отсутствие корреляции спектральных признаков;

2) равенство ковариационных матриц двух рассматриваемых классов;

3) проведение классификации объектов по тестовой выборке.

Разберем каждое из приведенных условий.

1. В предыдущей работе [5] обсуждалось понятие мерного эллипсоида рассеяния. С его помощью можно описать расстояние Махаланобиса между центрами двух нормальных распределений классов объектов, являющееся их мерой разделимости. При отсутствии корреляции спектральных признаков достигается максимальное значение объема эллипсоида, что свидетельствует о наибольшей разделимости двух классов и, как следствие, о наименьшей вероятности ошибки распознавания. Объем данного эллипсоида пропорционален определителю диагональной ковариационной матрицы одного из классов.

2. Нижняя граница вероятности ошибки распознавания двух нормально распределенных классов достигается при распределениях с равными ковариационными матрицами, верхняя – при равных средних значениях при условии распознавания по одному спектральному признаку [6].

3. Обучающие выборки, использованные для построения классификатора, могут быть затем сами им классифицированы. На практике классификация обучающего набора дает оптимистические результаты, т.е. вероятность ошибки, оцененная по обучающим данным, оказывается ниже вероятности ошибки для полного множества данных. Таким образом, можно сказать, что в лучшем случае классификация обучающих выборок дает нижнюю границу или оптимистический нижний предел истинной вероятности ошибки [7].

Метод определения объема обучающей выборки v

Из условий независимости спектральных признаков и равенства диагональных ковариационных матриц, описывающих многомерные нормальные распределения яркостей двух классов, следует равенство дисперсий яркостей относительно средних значений каждого спектрального признака для двух классов. Таким образом, отношение несмещенных оценок данных дисперсий будет подчиняться закону Фишера-Снедекора [8]:

$$F = \frac{S_{\max}^2}{S_{\min}^2} \sim \frac{\chi_v^2 / v}{\chi_f^2 / f},$$

где S_{\max}^2 и S_{\min}^2 – несмещенные оценки дисперсий яркостей относительно средних значений одного признака для каждого класса, χ_v^2 и χ_f^2 – случайные величины, имеющие распределение χ^2 с v и f – степенями свободы соответственно, $v = n_1 - 1$, $f = n_2 - 1$, где n_1 и n_2 – объем выборки для первого и второго класса соответственно.

По таблице распределения Фишера-Снедекора для заданного уровня значимости α определяется объем выборки.

Метод определения количества спектральных признаков n

В работе [5] описывался метод определения количества спектральных признаков для распознавания объектов с заданной точностью.

Для модельной ситуации распознавания двух классов с многомерными нормальными распределениями яркостных характеристик, равными ковариационными матрицами и равными априорными вероятностями появления данных классов справедливо выражение для ошибки распознавания в виде

$$P_e = \frac{1}{\sqrt{2\pi}} \int_{1/2\sqrt{J}}^{\infty} e^{-\frac{t^2}{2}} dt, \quad (1)$$

где J – дивергенция, или мера делимости двух классов по измеренным признакам в m спектральных каналах. Спектральные каналы соответствуют определенным длинам волн, на которых вычисляются признаки распознавания.

Для m -канальной спектральной аппаратуры информационное расхождение можно определить по формуле

$$J = m \frac{(B_1 - B_2)^2}{\sigma_1^2}, \quad (2)$$

где B_1 и B_2 – средние значения яркостей распознаваемых объектов в одном спектральном канале для первого и второго класса соответственно, а σ_1^2 – дисперсия яркостей относительно средних значений в одном спектральном канале. Приведенное соотношение справедливо при условиях: распределение яркостей ПО подчиняется нормальному закону; измеряемые признаки в разных каналах являются независимыми.

Задавшись вероятностью ошибки распознавания, можно определить дивергенцию двух классов объектов.

Вычисляя затем среднее значение $\mu = \frac{B_1 - B_2}{\sigma_1}$, нахо-

дят количество спектральных признаков n .

Заметим, что мы сокращаем количество спектральных каналов до n , поэтому количество спектральных признаков n меньше первоначального количества спектральных каналов m .

После того как определено количество спектральных признаков n , необходимо ответить на вопрос о значениях длин волн λ , соответствующих спектральным признакам распознавания.

Метод определения значений длин волн λ

Находя в каждом из спектральных каналов отноше-

ние разности матожиданий к дисперсии для двух классов можно вычислить все возможные комбинации, среднее которых будет равно определенному ранее значению μ . Воспользуемся t -критерием для одной выборки и проверим равенство матожиданий каждой из найденных комбинаций значению μ .

T -критерий для одной выборки позволяет проверить гипотезу о равенстве выборочного среднего некоторому заданному числу. В одновыборочных t -критериях, наблюдаемое среднее \bar{X} (вычисленное по реализации выборки отношений разности матожиданий к дисперсии для двух классов) сравнивается с ожидаемым (или эталонным) средним выборки μ (т.е. с некоторым теоретическим средним). Рассматриваются две гипотезы:

$$H_0: \bar{X} = \mu \text{ и } H_1: \bar{X} \neq \mu.$$

Статистика критерия $t = \frac{\bar{X} - \mu}{s / \sqrt{n}}$ имеет t -распре-

деление Стьюдента с $(n-1)$ степенью свободы, (в нашем случае n – количество найденных спектральных признаков). Выборочное стандартное отклонение s оценивается по наблюдаемой реализации выборки

$$s = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1}}.$$

Вычисленное значение t проверяют на предмет попадания в критическую область (критическое значение $t_{1-\frac{\alpha}{2}}(v)$ (где $v = n-1$) можно определить по таблицам).

Иными словами, необходимо проверить условие

$$|\bar{X} - \mu| > \frac{t_{1-\frac{\alpha}{2}}(v)}{\sqrt{n}} s.$$

Если неравенство не выполняется, гипотезу H_1 отвергают в пользу H_0 .

Приведенные известные положения математической статистики положены в основу разрабатываемого в настоящее время алгоритма классификации наземных объектов с тонкими спектральными различиями, в частности, здоровой пшеницы и пшеницы на ранних стадиях заражения бурой ржавчиной.

Математическое моделирование

Для математического моделирования использовались наземные ГС измерения двух классов объектов (здоровой и пораженной гербицидами пшеницы), полученные на Краснодарском тестовом участке в мае 2015 года с помощью спектрометра Ocean Optics MAYA 2000-Pro. Полевые измерения приводились к КСЯ, которые стали основой для статистической обработки спектральных характеристик пшеницы в программном пакете MSOffice Excel с помощью стандартных функций f -тест и t -тест, применение которых описано в работе. Последовательность измерений и предварительной обработки подробно описана в работе [9].

Для статистической обработки были выделены значения КСЯ в диапазоне от 400 до 1100 нм в 1637 спек-

тральных каналах. В электронных таблицах Excel были составлены матрицы КСЯ для первого и второго класса, и средствами пакета анализа данных был применен двухвыборочный Фишер-тест в соответствии с приведенными выше формулами, описанными в методе А, для проверки условия равенства ковариационных матриц по объему выборки для каждого класса. Уровень значимости α был выбран по умолчанию 0,05. Объем выборки был выбран одинаковым для обоих классов, т.е. $n_1 = n_2 = v$. В процессе математического моделирования в Excel для разного объема выборок определялось наименьшее количество спектральных каналов, в которых не выполняется условие равенства дисперсий для двух классов. В результате была построена зависимость объема выборки от количества спектральных каналов, которая представлена на рис. 1.

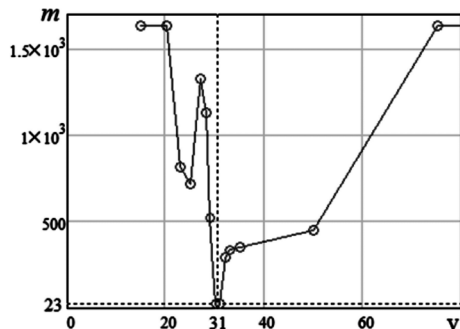


Рис. 1. Зависимость количества спектральных каналов m от объема выборки v (количества наблюдений), в которых не выполняется условие равенства дисперсий в спектральном канале для двух классов

Как видно из рисунка, наименьшее количество спектральных каналов, в которых не выполнено условие, равно 23. Такое значение соответствует количеству наблюдений 31.

Это количество наблюдений v , одинаковое для обоих классов, было взято за основу при создании двух обучающих выборок для проведения дальнейшего анализа и определения количества спектральных признаков n с учетом сделанных допущений 1 и 2.

Далее строился график (рис. 2) зависимости вероятности ошибки распознавания двух классов от дивергенции J в соответствии с соотношением (1). Задав вероятность ошибки P_e , можно определить $J_{зад}$ по графику и переписать формулу (2) в виде:

$$J_{зад} = n \frac{(B_1 - B_2)^2}{\sigma_1^2}.$$

Отсюда определяется количество спектральных признаков n для распознавания объектов с заданной вероятностью ошибки распознавания, предварительно вычисляя среднее отношение $\mu = \frac{B_1 - B_2}{\sigma_1}$ с учетом допущений 1, 2 и 3 для всех 1637 спектральных каналов. Для $P_e = 5\%$ количество спектральных признаков равно 6.

После найденного количества спектральных признаков n и объема обучающей выборки для каждого класса v анализировалась возможность определения длин

волн w , соответствующих найденным спектральным признакам, среднее отношение разности математических ожиданий к дисперсиям которых равно найденному значению

$$\mu = \frac{B_1 - B_2}{\sigma_1}.$$

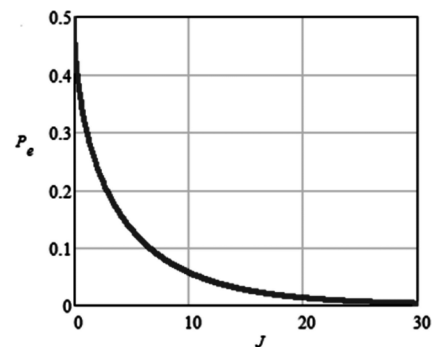


Рис. 2. Зависимость вероятности ошибки распознавания от дивергенции

Однако анализ показал, что количество комбинаций по 6 значений отношений разностей математических ожиданий для двух классов к дисперсии из 1637 составляет

$$C_{1637}^6 = \frac{1637!}{6!(1637-6)!} = 2.65 \cdot 10^{16},$$

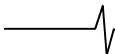
поэтому проверка на равенство среднему μ в соответствии с t -критерием не проводилась.

Обсуждение

Разработанный метод определения количества спектральных признаков n и соответствующих им значений длин w не целесообразен в применении для обработки информации с оптической аппаратуры с большим количеством спектральных каналов (порядка 1000). Однако такой метод может найти применение на первых этапах обработки космических ГС изображений, получаемых в меньшем количестве спектральных каналов (до 300 для космических ГС приборов), в связи с уменьшением количества комбинаций спектральных каналов, дающих необходимый набор спектральных признаков для распознавания двух классов объектов с заданной вероятностью ошибки, при сделанных допущениях D . Этот вопрос требует дальнейшей проработки.

Выводы и рекомендации

Благодаря проведенным исследованиям был разработан метод определения необходимого объема обучающих выборок в условиях оценки нижней границы вероятности ошибки статистического распознавания двух классов объектов с тонкими спектральными различиями, позволяющий улучшить процесс тематической обработки космических ГС данных. Улучшение процесса состоит в найденном подходе к определению минимально необходимого объема обучающих выборок v с учетом сделанных ограничений. На основе экспериментальных данных такой объем был определен и составил 31 для каждого класса объектов. Этот объем будет использован для получения уточненного объема выборок с учетом снятия наложенных условий, описанных в данной работе. Найденные выборки будут использованы в



дальнейшем для получения обучающих выборок при распознавании двух классов объектов на ГС изображениях, полученных из космоса. Поиску таких выборок и оценке вероятности ошибки распознавания двух классов объектов на основе найденных выборок будут посвящены дальнейшие исследования.

Работа выполняется при финансовой поддержке РФФИ, проект №16-44-230264 р_а.

Литература

1. Mader S., Vohland M., Jarmer T., Crop classification with hyperspectral data of the HyMAP sensor using different feature extraction techniques- Proceedings of the 2nd Workshop of the EARSeL SIG on Land Use and Land Cover. 28-30 September 2006.

2. Li C., Wang J., Wang L., Hu L., Gong P. Comparison of Classification Algorithms and training sample sizes in Urban Land Classification with Landsat Thematic Mapper Imagery –Remote Sensing. 2014. №6. P. 964-983.

3. Doma M.L., Gomaa M.S., Amer R.A. Sensitivity of pixel-based classifiers to training sample size in case of high resolution satellite imagery – Journal of Geomatics. 2015. V.9. № 1. P. 53-58.

4. Thenkabail Prasad S., Lyon G. John, Huete A. Hy-

perspectal Remote Sensing of Vegetation, CRC Press, USA, 2011, 782 p.

5. Третьяков В.А. Способ определения числа оптических каналов наблюдения природных объектов гиперспектральными средствами ДЗЗ для их классификации с заданной точностью, Тезисы докладов Второй международной научно-технической конференции «Актуальные проблемы создания космических систем дистанционного зондирования Земли». – М.:ОАО «Корпорация ВНИИЭМ», 2014г. С. 118-125.

6. Барабаш Ю.Л., Варский Б.В., Зиновьев В.Т., Кириченко В.С., Сапегин В.Ф. Вопросы статистической теории распознавания. – М.: Издательство «Советское радио», 1967. – 400 с.

7. Свейн Ф., Дейвис Ш. Дистанционное зондирование Земли: количественный подход. – М.: Издательство «Недра», 1983. – 401 с.

8. Андерсон Т. Введение в многомерный статистический анализ. – М.: Государственное издательство физико-математической литературы, 1963. – 500 с.

9. Акопов А.К., Баула Г.Г., Кривошеин В.В., Кротков А.Ю., Третьяков В.А. Разработка методики наземных валидационных измерений спектров сельскохозяйственных культур // Журнал «Космонавтика и ракетостроение». 2015. вып. № 6 (85). С. 45-50.

У в а ж а е м ы е а в т о р ы !

Редакция научно-технического журнала «Цифровая обработка сигналов» просит Вас соблюдать следующие требования к материалам, направляемым на публикацию:

1) Требования к текстовым материалам и сопроводительным документам:

- *Текст – текстовый редактор Microsoft Word.*
- *Таблицы и рисунки должны быть пронумерованы. На все рисунки, таблицы и библиографические данные указываются ссылки в тексте статьи.*
- *Объем статьи до 12 стр. (шрифт 12). Для заказных обзорных работ объем может быть увеличен до 20 стр.*
- *Название статьи на русском и английском языках.*
- *Рукопись статьи сопровождается: краткой аннотацией на русском и английском языках; номером УДК; сведениями об авторах (Ф.И.О., организация, должность, ученая степень, телефоны, электронная почта); ключевыми словами на русском и английском языках; актом экспертизы (при наличии в вашей организации экспертной комиссии).*

2) Требования к иллюстрациям:

Векторные (схемы, графики) – желательно использование графических редакторов Adobe Illustrator или Corel DRAW.

- *Растровые (фотографии, рисунки) – М 1:1, разрешение не менее 300dpi, формат tiff.*

УДК 004.932.4

АЛГОРИТМ КОМПЛЕКСИРОВАНИЯ ДВУХ РАЗНОСПЕКТРАЛЬНЫХ ИЗОБРАЖЕНИЙ НА ОСНОВЕ ЦВЕТОВОГО СИНТЕЗА

Шипко В.В., к.т.н., преподаватель кафедры автоматизации управления летательных аппаратов (и вычислительных систем) ВУНЦ ВВС «Военно-воздушная академия им. проф. Н.Е. Жуковского и Ю.А. Гагарина», e-mail: shipko.v@bk.ru.

THE ALGORITHM OF INTEGRATING TWO IMAGES OF DIFFERENT SPECTRUMS BASED ON COLOR SYNTHESIS

Shipko V.V.

Methods of displaying images of different spectrums have been analyzed. There is a new original algorithm of integrating two digital gray-scale images as a color image. This way we increase the contrast of the resulting image which contains elements of the original ones. These elements are gained from images of different spectral ranges. At the same time the algorithm facilitates the identification of spectrally dependent objects. There are some examples of integrated images presented in the article.

Key words: displaying images, spectrums, algorithm of integrating, contrast of the resulting image.

Ключевые слова: комплексирование изображений, многоспектральные изображения, цветные изображения, совместная обработка, цветовой синтез.

Введение

В последнее время активно развиваются комплексы многоканальной (многоспектральной, многозональной) съемки [1, 2]. Каждый из приемных каналов таких систем имеет свои преимущества и ограничения в применении в силу различных физических принципов формирования и представления двумерных изображений.

Изображения в оптическом диапазоне содержат информацию об отраженной или переотраженной энергии солнечного света, о химическом составе поверхностного слоя и т.д., в свою очередь изображения, полученные в инфракрасном диапазоне, несут информацию о распределении температурных полей, наконец, изображения, полученные в микроволновом диапазоне, предоставляют данные о геометрических и физических параметрах поверхности (шероховатость, диэлектрические свойства).

Существуют два обстоятельства, благодаря которым пред-почтительна одновременная съемка в нескольких спектральных диапазонах электромагнитного излучения [3]:

– априорная неопределенность спектральных отражательных (излучательных) свойств объекта и фона. Заранее не известно оптимальное расположение спектрального диапазона. Более того, ввиду многообразия фоно-объектовых ситуаций, расположение диапазона оказывается неоднозначным;

– в результате совместной обработки нескольких изображений одного и того же участка местности, полученных в различных спектральных диапазонах, можно дополнительно значительно повысить заметность объекта или его деталей. Это происходит благодаря возможности извлечения посредством такой обработки дополнительной информации об объектах местности, за-

Проведен анализ способов отображения многоспектральной видеоинформации. Предлагается оригинальный алгоритм комплексирования двух разноспектральных цифровых полутоновых изображений в виде цветного изображения, что позволяет повысить локальный контраст результирующего изображения, содержащего элементы исходных изображений одной и той же сцены, полученных в различных спектральных диапазонах, а также упростить выявление спектрально зависимых объектов. Представлены примеры комплексированных изображений.

ключенной в характере взаимосвязи зональных яркостей между спектральными диапазонами.

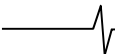
Анализ способов отображения многоспектральных данных

Существуют следующие подходы к отображению многоспектральной иконической информации [3, 4]:

- отдельно;
- синтез единого черно-белого изображения на основе линейной комбинации компонентов вектора зональных яркостей;
- синтез единого черно-белого изображения на основе нелинейных функций компонент вектора зональных яркостей;
- представление многозональной информации в виде единого цветного изображения в естественных или условных цветах.

Первый из подходов может использоваться в следующих случаях:

- когда анализ отдельно сформированных изображений в каналах достаточен для решения конкретной задачи;
- отсутствуют разработанные алгоритмы (либо знания, необходимые для конкретизации параметров алгоритмов) совместной визуализации измерений в разных каналах, повышающей эффективность анализа многоспектральных изображений;
- когда имеются вычислительные, либо другие трудности реализации таких алгоритмов.



Недостатками раздельного представления многоспектральных изображений являются:

- потеря части информации об изучаемой сцене, содержащейся в многоспектральных изображениях и связанной с корреляцией измерений в разных каналах;
- относительно большая продолжительность анализа в связи с тем, что приходится просматривать несколько, а не одно изображение, а также сопоставлять одно изображение с другим;
- игнорирование части зональных изображений (если нет возможности просматривать все изображения) может снизить эффективность дешифрирования многоспектральных изображений.

Достоинствами синтеза единого (комплексированного) изображения является:

- возможность аккумулирования в единое изображение особенностей спектросональных изображений объектов местности;
- возможность сделать доступной информацию многоспектральных изображений, связанную с корреляцией яркостей между каналами съемки, а точнее с различиями в корреляциях яркостей объектов и фонов.

Первое достоинство очевидно. Для иллюстрации второго рассмотрим возможные яркостные сечения двух гипотетических спектросональных изображений на рис. 1.

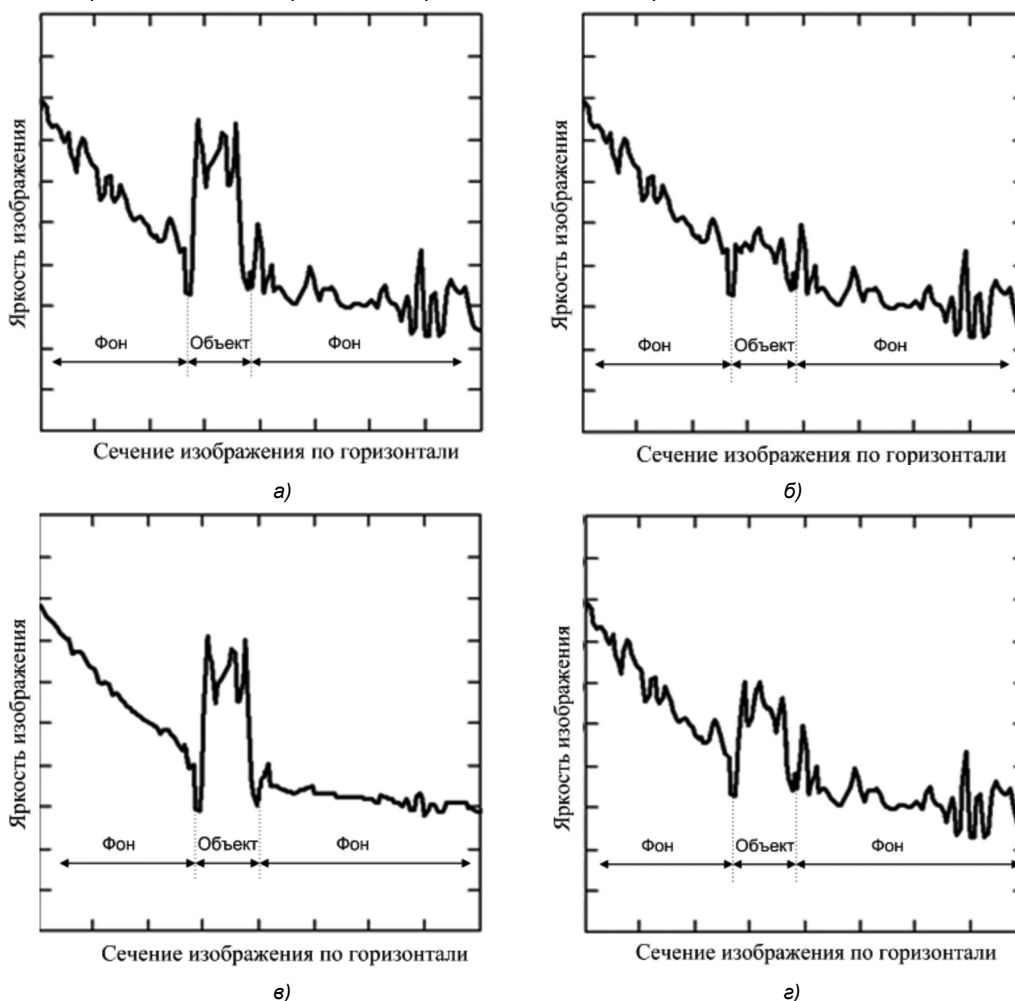


Рис. 1. Иллюстрация влияния корреляции яркостей между двумя спектросональными изображениями (а) и (б) на возможность повышения заметности объектов на едином комплексированном изображении при оптимальных (в) и неоптимальных (г) параметрах синтеза

Подобная ситуация часто реализуется при съемке в диапазонах отражательной области спектра излучения.

Недостатками синтеза единого комплексированного изображения является:

- необходимость предварительного выполнения затратной процедуры точного совмещения (взаимной привязки) спектросональных изображений;
- требуется определенная априорная информация для настройки параметров алгоритма синтеза;
- требуется некоторое время для выполнения собственно процедуры синтеза в цифровых системах обработки изображений.

Первый недостаток исключается при работе с изображениями от многоспектральных оптико-электронных систем, в которых совмещение происходит автоматически благодаря единой оптической системе. Третий частично компенсируется при использовании рекуррентных процедур вычислений в алгоритмах синтеза.

Предпосылки комплексирования разносспектральных изображений на основе цветового синтеза

Синтез цветного изображения используется с целью перевода различий классов в спектральной яркости в цветовые различия.

В простейшем случае синтез заключается в определении соответствия между номером спектрального изображения и номера компонента цветного изображения, например канал 1 – R (красная составляющая цветного единого комплексированного изображения), канал 2 – G (зеленая составляющая) и канал 3 – B (синяя составляющая).

В общем случае синтеза, исходное L – мерное яркостное пространство многоспектральных изображений отображается на трехмерное RGB-пространство. Другими словами, формируется три новых черно-белых изображения, каждое из которых – результат поточечного преобразования L исходных изображений. Подготовка полутонных RGB-составляющих цветного комплексированного изображения может производиться с использованием одного из методов синтеза единого черно-белого изображения.

Следует отметить, что практически все известные варианты синтеза цветного единого комплексированного изображения в той или иной мере имеют эвристический характер. Формализовать задачу оптимального получения цветосинтезированного изображения оказывается затруднительным. Это связано с относительно скудными данными по физиологии зрительного восприятия объектов на цветном и псевдоцветном изображениях.

Часто возникает вопрос о предпочтительности цветного или черно-белого синтеза. Он не имеет однозначного ответа. Все зависит от особенностей решаемой задачи дешифрирования. Однако очевидно, что цветовой синтез многоспектральных изображений является дополнительным эффективным инструментом, подчеркивающим спектрально-особенности объектов сцены.

Одним из частных вопросов является комплексирование разносспектральных изображений, полученных в достаточно разнесенных между собой диапазонах, например в видимом и инфракрасном диапазонах. Классическим подходом комплексирования таких изображений является получение единого полутонного изображения, включающего отличительные особенности исходных изображений [5, 6, 7]. Существуют алгоритмы формирования цветных комплексированных изображений, например, на основе методов преобразования цветов [8, 9], требующие качественных эталонов цвета и предварительных настроек. При этом все разработанные алгоритмы не позволяют качественно выделить спектрально зависимые объекты и не всегда результирующее комплексированное изображение получается с высоким локальным контрастом этих объектов. Это связано с ограниченностью динамического диапазона полутонных изображений и разнообразием фоновых ситуаций, при этом построение адаптивных алгоритмов весьма затруднительно.

Далее предлагается оригинальный алгоритм, на основе цветового синтеза, позволяющий повысить локальный контраст результирующего изображения, содержащего элементы исходных изображений одной и той же сцены, полученных в различных спектральных диапазонах, а также упростить выявление спектрально зависимых объектов.

Алгоритм комплексирования двух разносспектральных изображений

Формально алгоритм можно разбить на три основных этапа.

На первом этапе алгоритма осуществляется предварительная подготовка разносспектральных изображений (например, телевизионное (ТВ) и тепловизионное (ТПВ) изображение), далее определяют основное изображение с наибольшим количеством информативных деталей путем субъективной оценки изображений оператором, либо используя любой известный метод.

На втором этапе при комплексировании изображений вычисляют разности значений яркости каждого пиксела наиболее информативного изображения (например ТВ изображение) с окружающими его пикселями:

$$(\Delta_{i\pm p, j\pm q}^{\text{TB}})_d = (x_{i,j}^{\text{TB}} - x_{i\pm p, j\pm q}^{\text{TB}}), \quad (1)$$

где $i = 1, \dots, m; j = 1, \dots, n; m, n$ – число строк и столбцов ТВ изображения $x_{i,j}^{\text{TB}}$; $p = 0, \dots, P$ – параметр, определяющий координаты пиксела окрестности по i ; $q = 0, \dots, Q$ – параметр, определяющий координаты пиксела окрестности по j ; $d = 1, \dots, ((2P+1)(2Q+1)) - 1$ – индекс разности значений яркости i, j -го пиксела с пикселями окрестности.

Усредняют полученные разности по количеству пикселей окрестности с заданным коэффициентом усиления:

$$\Delta_{i,j}^{\text{TB}} = \left(\frac{\sum_d (\Delta_{i\pm p, j\pm q}^{\text{TB}})_d}{((2P+1)(2Q+1)) - 1} \right) k. \quad (2)$$

где k – коэффициент усиления средней разности яркостей пикселей наиболее информативного изображения.

Формируют комплексированное полутонное изображение путем суммирования значений яркости пикселей второго изображения (например, ТПВ изображение) с соответствующими значениями средней разности яркостей пикселей наиболее информативного изображения:

$$x_{i,j}^{\text{компл}} = x_{i,j}^{\text{ТПВ}} + \Delta_{i,j}^{\text{TB}}. \quad (3)$$

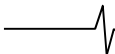
На третьем этапе, для визуализации отличительных признаков разносспектральных изображений, результирующее изображение представляют в виде синтезированного RGB-цветного изображения $\mathbf{X} = [\mathbf{R}_{i,j} \quad \mathbf{G}_{i,j} \quad \mathbf{B}_{i,j}]^T$

с тремя компонентами (например $\mathbf{R}_{i,j} = x_{i,j}^{\text{TB}}$, $\mathbf{G}_{i,j} = x_{i,j}^{\text{компл}}$, $\mathbf{B}_{i,j} = x_{i,j}^{\text{ТПВ}}$):

$$\mathbf{X}^{\text{компл}} = \begin{bmatrix} x_{i,j}^{\text{TB}} \\ x_{i,j}^{\text{компл}} \\ x_{i,j}^{\text{ТПВ}} \end{bmatrix}. \quad (4)$$

После чего выравнивают яркостный диапазон результирующего цветного изображения с помощью любого известного метода [10].

Блок-схема разработанного алгоритма представлена на рис. 2, где 1 – блок предварительной подготовки



изображений, выполняющий геометрические преобразование изображений, улучшение яркости и контраста, фильтрацию шума, 2 – блок выбора наиболее информативного изображения; 3 – блок вычисления разностей значений яркости каждого пикселя наиболее информативного изображения с окружающими его пикселями; 4 – блок вычисления среднего значения разностей яркости пикселей наиболее информативного изображения; 5 – блок усиления средних значений разностей яркости пикселей наиболее информативного изображения; 6 – блок формирования комплексированного полутонового изображения, согласно выражения (3); 7 – блок формирования результирующего RGB-цветного комплек-

сированного изображения, согласно выражения (4); 8 – блок выравнивания яркостного диапазона цветного изображения; 9 – блок вывода результирующего цветного изображения.

Пример комплексирования изображений

На рис. 3 (а, б) представлены исходные ТВ изображения (на рис. 3 (а) засветка объектов сцены, на рис. 3 (б) плохо различимые объекты с низким контрастом). На рис. 3 (в, г) представлены исходные ТПВ изображения с контрастными объектами. На рис. 3 (д, е) представлены результирующие комплексированные изображения полученные по одному из известных алгоритмов [5, 6]. На рис. 4

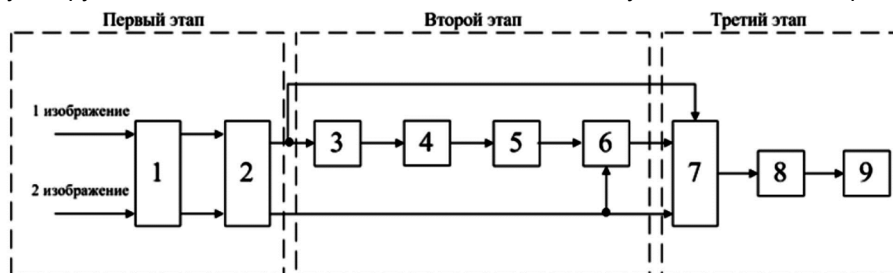


Рис. 2. Блок-схема алгоритма комплексирования двух разнеспектральных изображений



Рис. 3. Пример комплексирования ТВ изображений – а, б и ТПВ изображений – в, г по известным алгоритмам [5, 6] – д, е

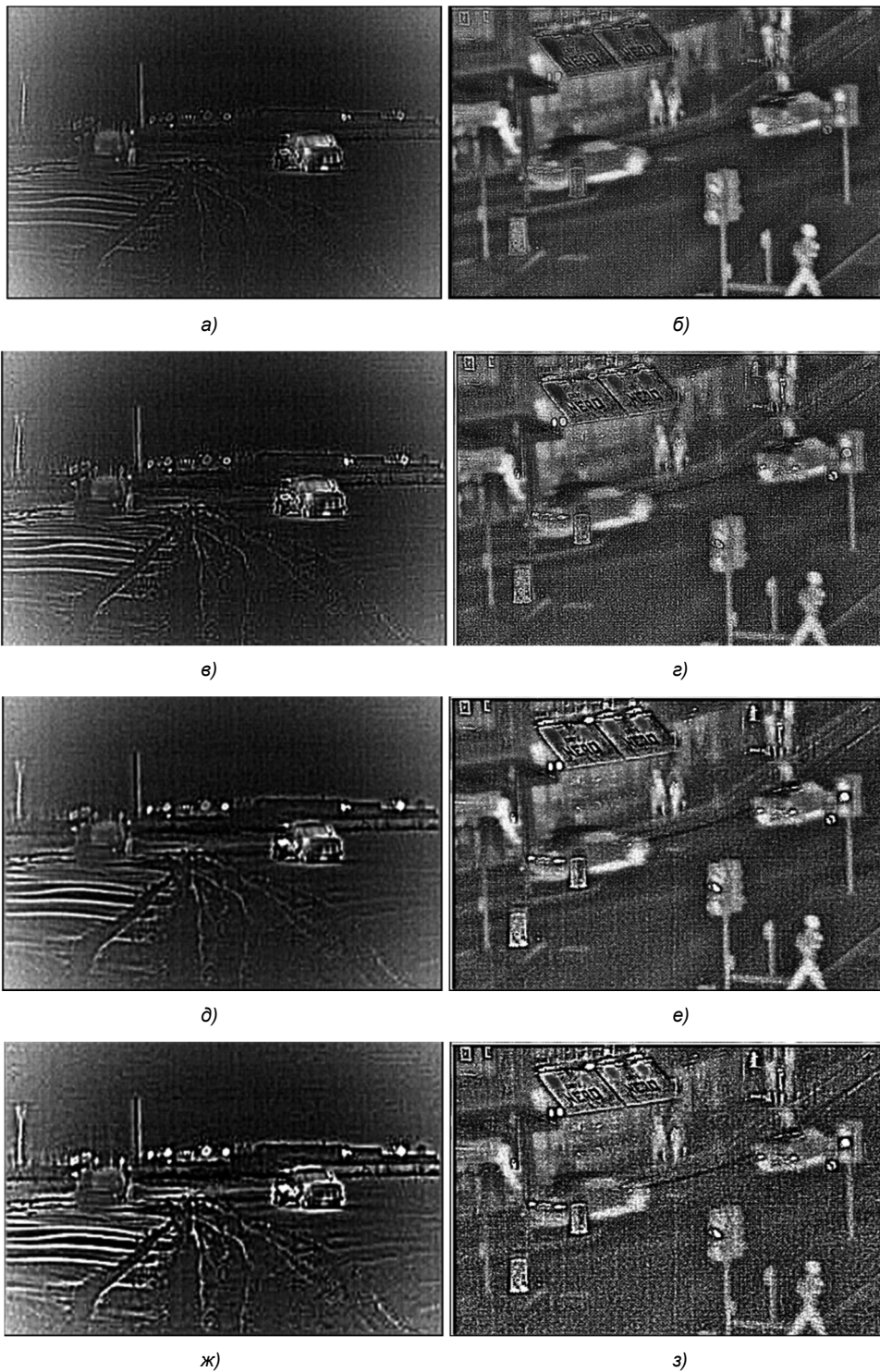


Рис. 4. Комплексированные полутоновые изображения полученные согласно выражения (3) для следующих исходных данных:

$(P, Q) = 1, k = 4$ – а, б; $(P, Q) = 1, k = 8$ – в, г; $(P, Q) = 3, k = 4$ – д, е; $(P, Q) = 3, k = 8$ – ж, з

представлены комплексированные полутоновые изображения, полученные на втором этапе предлагаемого алгоритма, согласно выражения (3). На рис. 5 представлены результирующие цветные комплексированные изображения, полученные на основе разработанного алгоритма, согласно выражения (4). Результирующее комплексированное цветное изображение является более информативным с высоким локальным контрастом и позволяет осуществлять селекцию спектрально зависимых объектов по их цветовому различию, как в условиях засветочных помех, так и в условиях низкого контраста

одного из исходных изображений.

Заключение

Таким образом, цветовой синтез является эффективным средством для комплексного дешифрирования изображений, полученных в различных областях электромагнитного спектра. Важным свойством синтеза является то, что результат синтеза по своим дешифровочным свойствам при прочих равных условиях не хуже, чем каждое из спектральнозональных изображений, что говорит в пользу его применения для сокращения времени де-



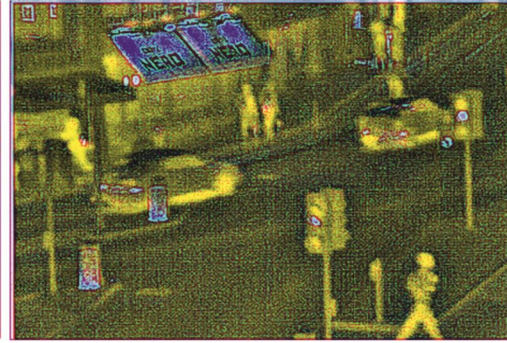
а)



б)



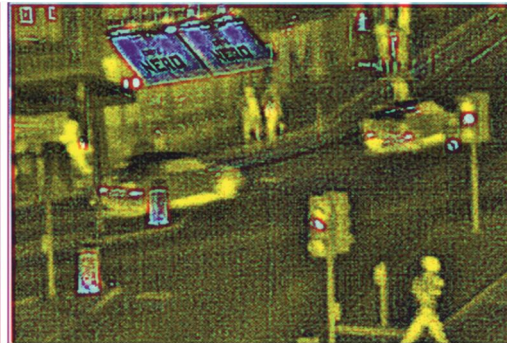
в)



г)



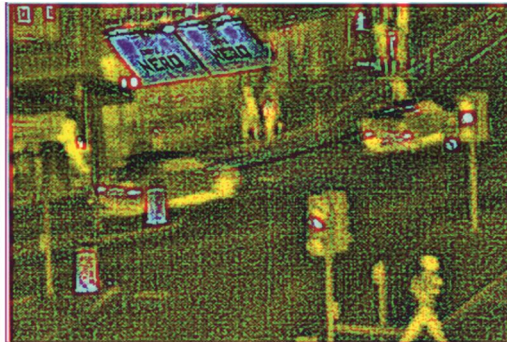
д)



е)



ж)



з)

Рис. 5. Комплексированные цветные изображения полученные согласно выражения (4) для следующих исходных данных:

$(P, Q) = 1, k = 4$ – а, б; $(P, Q) = 1, k = 8$ – в, г; $(P, Q) = 3, k = 4$ – д, е; $(P, Q) = 3, k = 8$ – ж, з

шифрования при исключении потерь информации. Его важной особенностью следует считать сохранение релевантной информации, необходимой для принятия решения. Цветовой синтез оказывается выгодным и с точки зрения вычислительных затрат, так как не требует обработки изображения попиксельно, а использует физиологи

ческие особенности цветового зрения человека. При этом вопрос автоматического обнаружения и распознавания объектов заданного интереса на многоспектральных изображениях остается открытым. Разработанный алгоритм может использоваться в любых многоспектральных следящих системах с целью повышения их информативности.

Литература

1. Тарасов В.В., Якушенков Ю.Г. Двух- и многодиапазонные оптико-электронные системы с матричными приемниками излучения. М.: Университетская книга; Логос. 2007. 192 с.
2. Действующие и перспективные спутники наблюдения земли // Геоматика. 2013. № 2. С. 107–111.
3. Автоматизированные системы наземных комплексов сбора и обработки данных воздушной разведки. Учебное пособие / Под ред. И.Н. Белоглазова. М.: ВВИА им. проф. Н.Е. Жуковского. 2003. 296 с.
4. Шипко В.В. О возможных подходах комплексирования иконической информации в многоканальных системах воздушно-космической разведки. Актуальные вопросы исследований в авионике: теория, обслуживание, разработки: сб. тезисов докл. IV Всероссийской научно-практической конференции «АВИАТОР», Воронеж, 16-17 февраля 2017 г., Воронеж: ВУНЦ ВВС «ВВА», 2017. 232 с.
5. Патент RU 2451338 опубликовано 20.05.2012 г., МПК G06T 5/00
6. Патент RU 2540778 опубликовано 10.02.2015 г., МПК G06T 5/50, G06K 9/46
7. Современные технологии обработки данных дистанционного зондирования Земли / Под ред. В.В. Еремеева. М.: Физматлит, 2015. 460 с.
8. Богданов А.П., Холопов И.С. Алгоритмы формирования цветного комплексированного изображения из многоспектральных монохромных на основе методов преобразования цветов / Цифровая обработка сигналов. 2013. № 3. С. 26-32.
9. Богданов А.П., Павлов О.В., Холопов И.С. Повышение быстродействия алгоритмов формирования цветного комплексированного изображения по сигналам разноспектральных монохромных датчиков / Цифровая обработка сигналов. 2013. № 3. С. 33-37.
10. Гонсалес Р., Вудс Р. Цифровая обработка изображений. М.: Техносфера, 2005. 1072 с.

Уважаемые коллеги!

Приглашаем Вас принять участие в формировании тематических выпусков журнала «Цифровая обработка сигналов» и размещению рекламы продукции (услуг) Вашей организации на его страницах. В случае положительного решения просим представить в редакцию журнала Ваши предложения по плановому размещению информационных материалов и макет рекламы продукции (услуг) с указанием желаемого её месторасположения: обложка (2-я, 3-я или 4-я стр.), цветная внутренняя полоса (объем полосы).

Журнал «Цифровая обработка сигналов» издается с 1999 года. Выходит ежеквартально, тиражом – 700 экз. Распространяется по подписке через агентство «Роспечать» в России (индекс 82185), СНГ и странах Балтии (индекс 20630), а также на Конференции: «Цифровая обработка сигналов и ее применение – DSPA'».

Научно-технический журнал «Цифровая обработка сигналов» включен в Перечень изданий, рекомендуемый ВАК РФ для публикации результатов научных исследований соискателями ученой степени доктора и кандидата технических наук в области радиотехники, связи, вычислительной техники, электроники, приборостроения, информационных технологий, информационно-измерительных и управляющих систем. По предварительным итогам за 2015 год по рейтингу Science Index базы РИНЦ (3,394) журнал «Цифровая обработка сигналов» занимает 344-ю позицию из почти 3000 представленных изданий. Импакт-фактор журнала за 5-летний период цитируемости составил 0,535!

Планируемые сроки издания отдельных номеров журнала:

- № 4 декабрь 2017 г. Тематический выпуск: «ЦОС в радиотехнике и системах телекоммуникаций».
- № 1 март 2018 г. Тематический выпуск: «ЦОС в радиотехнике и системах телекоммуникаций».
- № 2 июнь 2018 г. Тематический выпуск по материалам 20-й Международной научно-технической конференции «Цифровая обработка сигналов и ее применение-DSPA».
- № 3 сентябрь 2018 г. Тематический выпуск: «Цифровая обработка изображений».

Ориентировочная стоимость рекламных услуг:

- 4-я (внешняя) страница цветной обложки – 25 тысяч рублей.
- 2-я и 3-я (внутренние) страницы цветной обложки – 15 тысяч рублей.
- 1/2 цветной внутренней полосы – 8 тысяч рублей.

Ждем Ваших предложений.

С наилучшими пожеланиями, зам. главного редактора

д.т.н., профессор Витязев Владимир Викторович, телефон 8-903-834-81-81.

Предложения прошу направлять по адресу: E-mail: vityazev.v.v@rsreu.ru или info@dspa.ru

ИССЛЕДОВАНИЕ БЫСТРОДЕЙСТВИЯ АЛГОРИТМОВ ПОИСКА НА ИЗОБРАЖЕНИЯХ ОДНОИМЕННЫХ ТОЧЕК С ИСПОЛЬЗОВАНИЕМ СОВРЕМЕННОЙ ВЫЧИСЛИТЕЛЬНОЙ ТЕХНИКИ

*Кузнецов А.Е., зам. директора НИИ «Фотон» РГРТУ, д.т.н., профессор; foton@rsreu.ru;
Рыжиков А.С., инженер НИИ «Фотон» РГРТУ; foton@rsreu.ru.*

THE ANALYSIS OF THE SPEED OF SEARCH ALGORITHMS ON IMAGES OF THE CORRESPONDING POINTS USING MODERN COMPUTER

Kuznetsov A.E., Ryzhikov A.S.

The analysis of the speed of search algorithms on images of the corresponding points using modern computing techniques is conducted. The main stages of the work of the correlation-extreme algorithm and the algorithm based on SURF descriptors are considered. A theoretical and experimental evaluation of their performance is carried out using a multi-core CPU and video card. The recommendations on the choice of the finding reference points algorithm for the technology of automatic control of ERS materials geodesic reference are formulated.

Key words: points coregistration, correlation, SURF, graphics card.

Ключевые слова: идентификация одноименных точек, корреляция, SURF, видеокарта.

Введение

Базовым элементом технологии наземной обработки изображений от спутниковых систем высокого пространственного разрешения является контроль геодезической привязки маршрутов съемки. Эта операция выполняется путем распознавания на анализируемом и опорном изображениях одноименных объектов и измерения их координат [1, 2].

Известны два подхода к поиску одноименных точек на паре снимков. Первый, традиционный, основан на корреляционном сопоставлении фрагментов анализируемого снимка с множеством его образов в зоне поиска другого. Второй подход предполагает сопоставление дескрипторов локальных особенностей анализируемого и опорного изображений. Поскольку маршрут съемки представляет собой массив данных сверхбольшого объема (десятки и сотни гигабайт), то предъявляются высокие требования по скорости выполнения операции отождествления на снимках одноименных объектов. В этой связи возникает задача исследования быстродействия рассмотренных алгоритмов на современной вычислительной технике, включающей многоядерный центральный процессор (ЦП, CPU) и видеокарту (GPU). Результатом такого исследования будут рекомендации по выбору алгоритма поиска опорных точек и его практической реализации в технологии автоматического контроля геопривязки материалов съемки.

Архитектура современной ЭВМ

Современный ЦП состоит из нескольких ядер (1 – 36), оптимизированных для обработки последовательных инструкций с максимальной производительностью.

Выполнено исследование быстродействия алгоритмов поиска на изображениях одноименных точек с использованием современной вычислительной техники. Рассмотрены основные этапы работы корреляционно-экстремального алгоритма и алгоритма на основе SURF-дескрипторов. Проведена теоретическая и экспериментальная оценка их производительности при реализации на многоядерном центральном процессоре и видеокarte. Сформулированы рекомендации по выбору алгоритма поиска опорных точек для технологии автоматического контроля геодезической привязки материалов ДЗЗ.

GPU строится на основе масштабируемого массива потоковых мультипроцессоров и имеет тысячи упрощенных и энергоэффективных ядер, оптимизированных для массивно-параллельной обработки данных. Блоки векторных вычислений ЦП реализуют принцип SIMD (одна инструкция выполняется над многочисленными данными), а GPU – SIMT (одна инструкция и несколько потоков). SIMT не требует векторизации данных и допускает произвольные ветвления в потоках. Упрощенная архитектура ЭВМ приведена на рис. 1.

Скорость доступа к данным в памяти GPU превосходит или соответствует скорости доступа к ОП, а чтение/запись разделяемой памяти (Shared Mem) на несколько порядков быстрее. Выполнение расчетов на GPU эффективно, когда одну и ту же последовательность математических операций применяют к большому объему данных. При этом лучшие результаты достигаются, если отношение числа арифметических инструкций к числу обращений к памяти достаточно велико. Это предъявляет меньшие требования к управлению исполнением, вырожденному в GPU, а высокая плотность арифметических инструкций и большой объем входных данных уменьшает ценность больших кэшей, присутствующих в CPU.

С учетом этих особенностей, исследование быстродействия анализируемых алгоритмов будем выполнять применительно к их реализации на CPU и GPU.

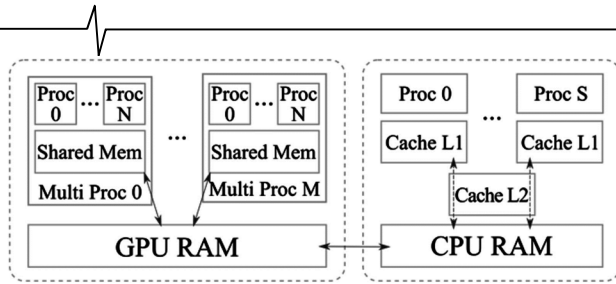


Рис. 1. Архитектура современной ЭВМ

Алгоритмы поиска одноименных точек

Алгоритм пирамидального корреляционно-экстремального (КЭ) поиска и алгоритм, основанный на дескрипторах SURF (далее алгоритм SURF), подробно рассмотрены в работах [2, 3]. Здесь приведем лишь краткое их описание.

Алгоритм поиска одноименных точек на основе корреляционного отождествления состоит из следующих шагов:

Формирование множества координат точек $C = \{c_i\}$, $c_i = \{x_i, y_i\}$ анализируемого изображения $b(x, y)$, для которых необходимо найти одноименные на опорном изображении $b'(x', y')$.

Определение для каждой точки c_i зоны поиска – множества возможных координат одноименной точки в системе координат опорного изображения $C'_i = \{c'_{it}\}$, $c'_{it} = \{x'_{it}, y'_{it}\}$.

Принятие решения о отождествлении всех пар (c_i, c'_{kt}) , для которых выполняется условие:

$$\begin{cases} k = \arg \max_i r(c_i, c'_{it}) \\ r(c_i, c'_{kt}) > r_{\min} \end{cases}$$

где $r(c_i, c'_{it})$ – функция, определяющая нормированную корреляцию фрагмента анализируемого и опорного изображения с центрами в точках с координатами c_i и c'_{it} соответственно; r_{\min} – пороговое значение корреляционной функции.

В работах [3, 4], предложены подходы, позволяющие снизить вычислительную сложность и повысить надежность идентификации одноименных точек в условиях значительных координатных рассогласований и сюжетных различий анализируемого и опорного изображения. Эти оптимизации прямо влияют на способ организации вычислений.

Число операций сопоставления фрагментов при корреляционно-экстремальном отождествлении составляет $T_1 = (M'-M+1) \cdot (N'-N+1)$, где M' – высота фрагмента, N' – ширина, M' и N' – высота и ширина зоны поиска соответственно. Квадратичный рост числа операций с увеличением зоны поиска ограничивает возможность непосредственного применения алгоритма при больших взаимных рассогласованиях снимков. Для повышения быстродействия процесса идентификации образа фрагмента в большой зоне поиска применяется механизм пирамидального представления изображений, позволяющий осуществить последовательное уточнение начальных приближений. Пирамида представляет собой последовательность изображений, в которой каждое последующее изображение является прорежен-

ной вдвое копией предыдущего. Очевидно, что на i -ом уровне пирамиды область поиска сокращается в $2^{2(i-1)}$ раз. Число операции сопоставления фрагментов соответственно составит $T_2 = \sum_i \frac{T_1}{2^{2(i-1)}}$. В работе [3] отме-

чено, что оптимальное число уровней пирамиды 2 – 3. Увеличение числа уровней не имеет смысла, т.к. при каждом прореживании происходит снижение информативности изображений, что повышает вероятность попадания в ложный локальный экстремум корреляционной функции при поиске образа загрубленного фрагмента.

Рассмотрим теперь основные шаги алгоритма SURF. Обозначим $\tau(b, c_i)$ – критерий, определяющий, является ли точка c_i характерной точкой изображения $b(x, y)$. Дескриптором G_{c_i} точки c_i назовем вектор чисел, характеризующий распределение интенсивностей пикселей фрагмента с центром в точке c_i . В качестве меры близости дескрипторов точек c_1, c_2 применяется Евклидово расстояние $d(G_{c_1}, G_{c_2})$.

Алгоритм поиска одноименных точек на основе дескрипторов SURF состоит из следующих шагов:

Формирование множества точек $C = \{c_i\}$ анализируемого изображения $b(x, y)$, для которых выполняется $\tau(b, c_i)$, и множества соответствующих дескрипторов $F = \{G_{c_i}\}$.

Формирование множества точек $C' = \{c'_i\}$ опорного изображения $b'(x', y')$, для которых выполняется $\tau(b', c'_i)$, и множества соответствующих дескрипторов $F' = \{G_{c'_i}\}$.

Формирование множества пар одноименных точек $O = \{(c_i, c'_i)\}$, для которых выполняется условие

$$\begin{cases} \forall G_{c_k} \in F \rightarrow d(G_{c_k}, G_{c'_i}) > d(G_{c_i}, G_{c'_i}) \\ \forall G_{c'_k} \in F' \rightarrow d(G_{c_i}, G_{c'_k}) > d(G_{c_i}, G_{c'_i}) \end{cases}$$

Авторами SURF предложен ряд оптимизаций вычислительного процесса [5].

При детектировании характерных точек для вычисления элементов матрицы Гессе и фильтра Хаара используется интегральное представление изображения, по которому можно быстро найти сумму яркостей пикселей в произвольной прямоугольной области изображения.

Традиционно элементы матрицы Гессе вычисляются путем свертки фрагмента изображения с фильтрами, изображенные на рис. 2 а. В детекторе SURF применяется аппроксимация лапласиана гауссиан (Fast-Hessian), рассчитываемая с использованием фильтров, изображенных на рис. 2 б. Такие свертки эффективно вычисляются с привлечением интегрального изображения.

Для быстрого сопоставления дескрипторов анализируемого и опорного изображения строится древовидная структура, как показано на рис. 3. Далее производится вычисление меры близости только для пар дескрипторов, обладающих одинаковым значениями знака Гессиа-на, масштаба и сходными направлениями градиента.

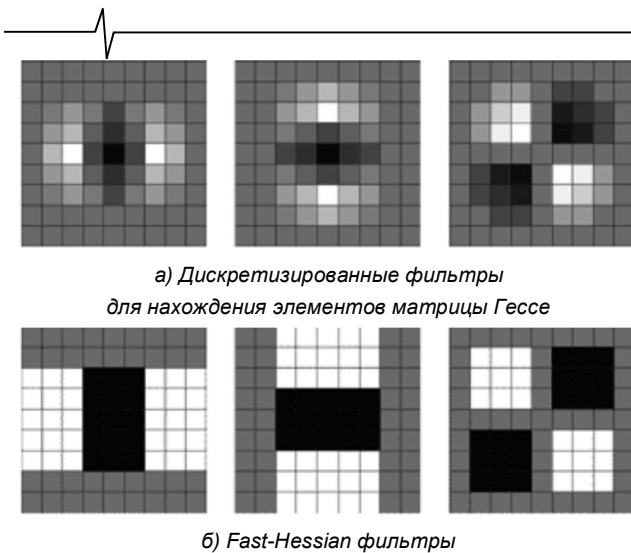


Рис. 2. Фильтры для вычисления лапласиана гауссиан

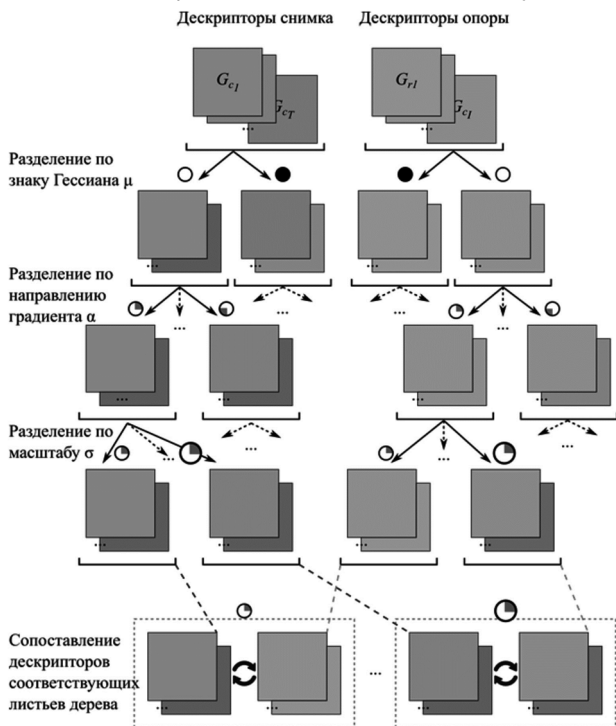


Рис. 3. Дерево сопоставления дескрипторов

Оценка быстродействия алгоритмов поиска одноименных точек

Выполним оценку количества тактов процессора Intel с архитектурой Skylake при однопоточном выполнении векторизованного (AVX) кода, необходимых для поиска пары одноименных точек корреляционно-экстремальным отождествлением и с использованием SURF-дескрипторов. Примем допущение об отсутствии кэш-промахов и накладных расходов на векторизацию данных. Процесс поиска характерных точек в обоих случаях рассматривать не будем.

Пусть размер фрагмента составляет 128 пикселей, допустимое рассогласование координат 512 пикселей. Число операций сопоставления фрагментов при наличии трех уровней пирамиды приведено в табл. 1 Таблица (принято допущение, что ошибка определения положения фрагмента на i -ом уровне составляет не более 2 пиксе-

лей на $i-1$ -ом уровне). Число арифметических операций для вычисления корреляционной функции фрагментов размером $M \times N$ приведено в табл. 2. (Опущено вычисление средней яркости, т.к. эту операцию можно свести к расчету среднего значения одной строки/столбца для каждого положения фрагмента). В соответствии с приведенными данными, для обнаружения образа фрагмента необходимо по меньшей мере $5 \cdot (32^2 \cdot 9216 + 64^2 \cdot 64 + 128^2 \cdot 64) \approx 5,3 \cdot 10^7$ операций сложения и $3 \cdot (32^2 \cdot 9216 + 64^2 \cdot 64 + 128^2 \cdot 64) \approx 3,2 \cdot 10^7$ операций умножения целых чисел. В табл. 3 приведено число тактов, за которое выполняется каждая из используемых операций. Общее число тактов для обнаружения образа фрагмента составит не менее $4,3 \cdot 10^7$.

Таблица 1. Параметры пирамидального поиска при КЭ отождествлении

Уровень пирамиды	Размер зоны поиска	Размер фрагмента	Число операций сопоставления фрагментов
3	128x128	32x32	9216
2	72x72	64x64	64
1	136x136	128x128	64

Таблица 2. Число арифметических операций, необходимых для расчета корреляционной функции

Операция	Число операций
+/-	$5MN$
*	$3MN$
$\sqrt{\quad}$	1

Таблица 3. Производительность векторизованных команд архитектуры Skylake

Операция	Число тактов (AVX)
+/-	0,5
*	0,5
$\sqrt{\quad}$	12

Оценим теперь число операций, необходимых для построения дескрипторов и вычисления меры близости между ними.

Дескриптор формируется путем расчета 25 фильтров Хаара в 16 квадрантах. Для расчета фильтра Хаара с использованием интегрального изображения необходимо 12 операций сложения. Таким образом, для построения дескриптора требуется произвести 4800 операций сложения при вычислении фильтров Хаара и 960 операций сложения при обобщении точечных градиентов, всего $4800 + 960 \approx 5,7 \cdot 10^3$ операций сложения, $2,9 \cdot 10^3$ тактов. Вычисление меры близости пары дескрипторов (Евклидово расстояние между 64-компонентными векторами) предполагает 127 операций сложения, 64 умножения и одно извлечение квадратного корня, 121 такт.

При выделении четырех уровней масштаба и четырех секторов направления градиента (рис. 2), дескрипторы анализируемых изображений в лучшем случае будут разбиты на 32 непересекающиеся группы равных размеров. Число операций сопоставления A дескрипто-

ров одного изображения с тем же числом дескрипторов другого изображения составит $32\left(\frac{A}{32}\right)^2 = \frac{A^2}{32}$. В худшем случае, если все признаки, по которым осуществляется выделение групп, вырождены, сопоставление полным перебором потребует A^2 операций.

Примем среднее число операций сопоставления дескрипторов за $\left(\frac{A^2}{32} + A^2\right)/2 = \frac{A^2}{16}$. Тогда число тактов для формирования A дескрипторов каждого изображения и их сопоставления составит $2 \cdot A \cdot (2,9 \cdot 10^3) + A^2/16 \cdot 121$.

Согласно приведенным расчетам, время выполнения КЭ поиска одного фрагмента эквивалентно времени создания и сопоставления $2 \cdot 10^3$ пар дескрипторов характерных точек. Ожидаемая производительность рассмотренных алгоритмов приведена на рис. 4 (ось числа тактов в логарифмическом масштабе).

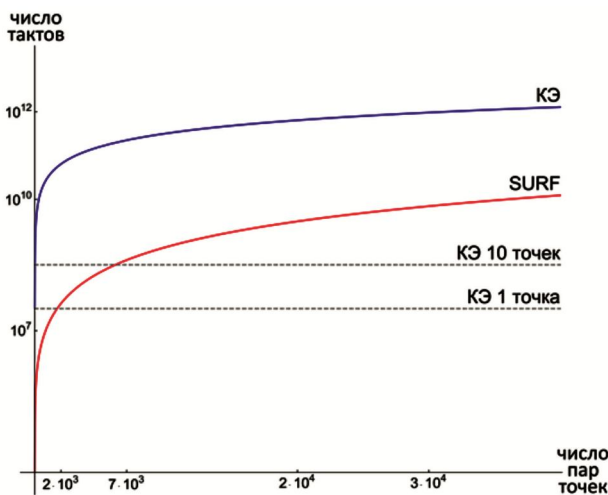


Рис. 4. График теоретической производительности рассмотренных алгоритмов в однопоточной реализации

Полученные оценки быстродействия соответствуют однопоточному выполнению алгоритмов поиска одноименных точек на центральном процессоре. Однако средства современной вычислительной техники обладают многоядерным центральным процессором и графическим сопроцессором (видеокартой). Точная теоретическая оценка быстродействия алгоритмов на таких устройствах практически невозможна в связи со сложным конкурентным доступом к разделяемым данным и многоуровневой иерархией кэш-памяти.

Для получения экспериментальных данных по быстродействию выполнена многопоточная реализация алгоритмов на языке C++ с использованием примитивов OpenCV. Производительность реализованных алгоритмов тестировалась на конфигурации CPU Intel Xeon 2.4 Ghz x 36, GPU NVIDIA Tesla k20c (706 MHz x 2496).

Для КЭ алгоритма в каждом потоке параллельно выполняется построение пирамидального представления эталона и фрагмента зоны поиска и КЭ поиск. График, иллюстрирующий быстродействие многопоточной реализации в зависимости от числа потоков, приведен на рис. 5.

Из графика следует, что увеличение числа задействованных для вычисления ядер центрального процессора не приводит к пропорциональному ускорению. Алгоритм КЭ отождествления масштабируется слабо. Вероятно, это связано с тем, что время, затрачиваемое на обращение к данным в ОП, преобладает над затратами на собственно вычисления. С ростом числа одновременно выполняющихся потоков эта особенность приводит к большому числу кэш-промахов и не позволяет задействовать все ресурсы CPU.

Изучена возможность реализации алгоритма КЭ отождествления на GPU. Проведенные эксперименты выявили принципиальную проблему. Для определения истинного положения фрагмента необходимо на завершающем этапе выполнить поиск максимального / минимального значения корреляционной функции, что сводится к последовательному проходу по всем пикселям раstra, определяющего величину КЭ функции в каждой точке зоны поиска. Проведение этой последовательной по своей природе операции «на месте», в GPU, крайне неэффективно, т.к. при этом простаивает большая часть ядер. Пересылка раstra на CPU и дальнейший поиск экстремумов на ресурсах ЦП дала прирост скорости 40%. Тем не менее это на порядок более медленная, чем вычисление КЭ функции в памяти GPU, операция определяет верхнюю границу скорости обработки данных. Анализ реализованной программы показал, что 81% времени уходит на пересылку данных CPU ↔ GPU, а быстродействие в целом эквивалентно реализации КЭ поиска в 8 потоков на CPU.

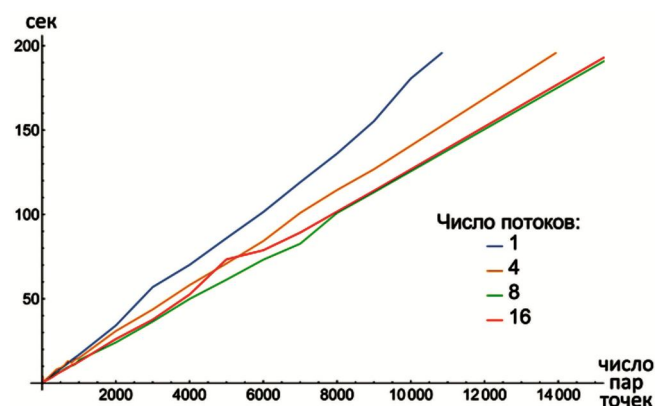


Рис. 5. График быстродействия многопоточной реализации КЭ алгоритма

В многопоточной реализации алгоритма SURF на центральном процессоре построение пирамиды разностей гессиан производится последовательно, по слоям. Параллельно выполняется детектирование характерных объектов и построение их дескрипторов, при этом удается задействовать все 36 ядер процессора. Сопоставление дескрипторов из соответствующих листьев дерева каждого изображения так же проводится в отдельных потоках с эффективным использованием 8 – 16 ядер процессора в зависимости от структуры дерева дескрипторов.

В отличие от GPU-версии алгоритма на графическом процессоре организовано параллельное вычисление гессиан и их сглаживание, детектирование характерных объектов и построение дескрипторов. Сопоставление дескрипторов выполняется параллельно на центральном процес-

соре, как и в CPU реализации. Результаты выполнения многопоточной реализации алгоритма SURF и КЭ (ЦП, 8 потоков) приведены на рис. 6 (ось времени представлена в логарифмическом масштабе). Столь высокой производительности алгоритма SURF удалось добиться в первую очередь за счет высокой локальности и малого объема данных, необходимого для формирования дескриптора. Это позволяет естественным образом параллельно рассчитывать большое их количество с использованием быстродействующей кэш памяти процессора (CPU-версия) или разделяемой памяти блоков (GPU).

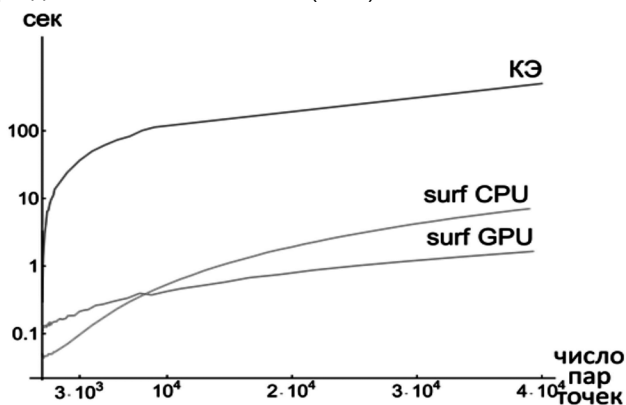


Рис. 6. График быстродействия многопоточной CPU и GPU реализации рассмотренных алгоритмов

Таблица 4. Число точек, сопоставленных за фиксированное время

Длительность, с	Число сопоставленных пар точек		
	КЭ, 8 ядер	SURF, CPU	SURF, GPU
1	68	14170	24940
2	159	20400	45700
3	198	25500	60000

Вывод: соотношение производительности многопоточных реализаций алгоритмов КЭ и SURF сильно коррелирует с теоретически рассчитанными величинами для однопоточных вариантов. Многопоточный алгоритм SURF, реализованный на GPU, производительнее многопоточной CPU-версии более чем в 2 раза. Число пар точек, сопо-

ставленных за фиксированное время в 1, 2 и 3 секунды реализованными алгоритмами, приведено в табл. 4.

Заключение

Алгоритм, основанный на дескрипторах SURF, может быть эффективно распараллелен на современной вычислительной технике. Его массивно-параллельная реализация как на многоядерном центральном процессоре, так и с привлечением ресурсов видеокарты, позволяет обеспечить формирование массивов координат одноименных точек со скоростью, на два порядка превышающей скорость корреляционно-экстремального поиска, так же реализованного параллельно. Другой проблемой, определяющей возможность практического применения алгоритма SURF, остается вопрос надежности сопоставления. Это требует дополнительного исследования, которому планируется посвятить следующую работу.

Литература

1. Кузнецов А.Е. Конвейерные технологии потоковой обработки данных от высокодетальных систем ДЗЗ// Тез. докл. четвертой международной науч.-техн. конф. «Актуальные проблемы создания космических систем дистанционного зондирования земли» – М.:АО «Корпорация «ВНИИЭМ», 2016. – 220 с.
2. Кузнецов А.Е., Пошехонов В.И., Рыжиков А.С. Технология автоматического контроля точности геопривязки спутниковых изображений по опорным снимкам от КА «Landsat-8» // Цифровая обработка сигналов. – 2015. – № 3. – С. 37-42.
3. Кузнецов А.Е., Светелкин П.Н. Оперативное совмещение спектральнональных изображений при цветосинтезе // Вестник Рязанского государственного радиотехнического университета. Выпуск 22, 2007. С. 3-7.
4. Прэтт У. Цифровая обработка изображений. В 2-х кн.: пер. с англ. М.: Мир, 1982. Кн.1: 312 с., кн.2: 480 с.
5. H. Bay, T. Tuytelaars, L. Van Gool, Surf: Speeded up robust features // Computer vision—ECCV. – 2006. – С. 404-417.

НОВЫЕ КНИГИ



Витязев В.В. Многоскоростная обработка сигналов – М.: Изд-во «Горячая линия-Телеком», 2017 г. – 336 с.: ил.

Рассмотрена эволюция теории и технологий многоскоростной обработки сигналов в период с начала 70-х гг. прошлого столетия до наших дней с позиции вклада, который внесли в их развитие работы российских ученых и специалистов в области цифровых информационных технологий реального времени. Описаны методы и алгоритмы многоступенчатой и многокаскадной реализаций цифровых узкополосных фильтров и банков цифровых фильтров на основе эффектов прореживания по времени (децимация во временной области) и по частоте (децимация в частотной области).

Приведена методика оптимизации параметров многоступенчатых и многокаскадных структур цифровых полосовых фильтров. Построение оптимальных структур и расчет параметров фильтров частотной селекции иллюстрируется многочисленными примерами.

Для специалистов, научных работников, преподавателей вузов, аспирантов; будет полезна студентам информационных и инфокоммуникационных направлений подготовки.

УДК 621.391

РАЗРАБОТКА И ИССЛЕДОВАНИЕ АЛГОРИТМОВ ОБРАБОТКИ И РАСПОЗНАВАНИЯ РЕЧЕВЫХ СИГНАЛОВ И ИЗОБРАЖЕНИЙ ДЛЯ СИСТЕМ МУЛЬТИМОДАЛЬНОЙ БИОМЕТРИИ

Хрящев В.В., к.т.н., доцент Ярославского государственного университета им. П.Г. Демидова, e-mail: vhr@yandex.ru;

Приоров А.Л., д.т.н., доцент Ярославского государственного университета им. П.Г. Демидова, e-mail: andcat@yandex.ru;

Стефаниди А.Ф., аспирант Ярославского государственного университета им. П.Г. Демидова, e-mail: antonstefanidi@mail.ru;

Топников А.И., к.т.н., доцент Ярославского государственного университета им. П.Г. Демидова, e-mail: topartgroup@gmail.com.

DEVELOPMENT AND ANALYSIS OF SPEECH AND IMAGE SIGNALS PROCESSING FOR MULTIMODAL BIOMETRICS SYSTEMS

Khryashchev V.V., Priorov A.L., Stefanidi A.F., Topnikov A.I.

The paper considers bimodal personality recognition problem using audio and video data by analyzing the speaker face and voice. Two speaker identify algorithms are proposed and compared in this paper. The work of the first algorithm consists of feature extracting from the speech signal in the form of mel-frequency cepstral coefficients and forming on their basis a speaker model using Gaussian mixtures, the second is based on the use of a universal background model obtained from the records of a large number of speakers.

The simulation results demonstrated the possibility of applying the proposed algorithms in the person identification problem with the help of combining biometric features.

Key words: digital speech processing, digital image processing, machine learning, speaker identification, Gaussian mixture models, face recognition, convolutional neural network, bimodal biometrics.

Ключевые слова: цифровая обработка речевых сигналов, цифровая обработка изображений, машинное обучение, идентификация диктора, модели гауссовых смесей, распознавание лиц, сверточная нейронная сеть, бимодальная биометрия.

Введение

Целью работы является разработка алгоритмов бимодального распознавания личности с использованием аудио- и видеоданных путем анализа лица и голоса диктора.

Такой подход имеет ряд преимуществ относительно систем унимодальной биометрии [1]. Во-первых, он позволяет повысить уровень надежности системы и усложняет возможность фальсификации данных. Во-вторых, распознавание по лицу и голосу дает возможность получения биометрических параметров в отсутствие физического контакта человека с системой (неинвазивная система), что расширяет спектр практического использования предлагаемой технологии. В-третьих, это повышает точность и устойчивость работы системы биометрической идентификации.

В ходе исследования решались следующие задачи: реализация текстового словаря; сбор, накопление и структурирование собственной базы лиц и речевых сигналов; реализация и сравнение методов идентификации диктора с применением гауссовых смесей; моделирова-

Рассматривается задача бимодального распознавания личности с использованием аудио- и видеоданных путем анализа лица и голоса диктора. Предложены два алгоритма идентификации диктора и произведено их сравнение. Работа первого алгоритма основывается на извлечении из речевого сигнала признаков в виде мел-частотных кепстральных коэффициентов и формирования на их основе модели диктора с применением гауссовых смесей. Работа второго основана на применении универсальной фоновой модели, полученной на основе записей большого числа дикторов. Результаты моделирования продемонстрировали возможность применения предложенных алгоритмов в задаче идентификации личности путем комбинирования биометрических признаков.

ние и анализ работы алгоритма распознавания лиц на основе сверточной нейронной сети; анализ бимодальной системы распознавания личности по голосу и лицу.

Исследование алгоритмов идентификации диктора Подготовка текстового словаря и формирование базы речевых сигналов

Для анализа работы алгоритмов обработки и распознавания речевых сигналов сформирована собственная база на 100 классов. Каждый класс представлял собой конкретного человека – диктора, которому в базе соответствовало 5 аудиозаписей: 3 для обучения и 2 тестовых. В сумме общий объем экспериментальных аудиоданных составлял 500 речевых сигналов. Для записи голосов были за-

действованы студенты физического факультета ЯрГУ им. П.Г. Демидова. Средний возраст дикторов 20–25 лет.

В качестве информационного наполнения аудиозаписей предложено реализовать тестовый словарь типа «число-слово-число-слово-число» (ЧСЧСЧ). Сформированный словарь собран в виде текстового документа, содержащего 500 последовательно идущих и неповторяющихся строк. Каждая звуковая дорожка представляет собой запись голоса диктора, который читает одну строчку ЧСЧСЧ. Все числа при записи проговаривались в виде набора цифр. Например, первая строка интерпретировалась диктором как: «четыре, восемь, три, прохождение, два, один, четыре, поселенка, четыре, пять, три, четыре». Таким способом первый диктор проговаривал по отдельности пять последовательно идущих строк. Строки с шестой по десятую использовались для записи речи второго диктора. И так для каждого из 100 классов. В табл. 1 представлен фрагмент словаря.

Реализация словаря выполнена с применением среды разработки MATLAB. В процессе моделирования использовался лексикографический словарь А.А. Зализняка, состоящий из 93392 слов [2]. При генерировании чисел и слов применялась встроенная функция генерации случайных чисел «randit». На модель накладывались следующие ограничения: слова должны содержать не менее 7 букв, первое и второе число должны быть в диапазоне [100:999], а третье число – [1000:9999]. Для уменьшения вычислительной сложности задачи использовались следующие технические характеристики аудиозаписи: формат записи WAV; частота дискретизации 8 кГц; разрядность квантования 16 бит; скорость потока 128 кбит/с; одноканальная (моно) звукозапись сигнала. Сбор аудиоданных осуществлялся с применением мобильного устройства SAMSUNG GT-i9260 и приложения Smart Recorder by SmartMob v. 1.8.0.

Анализ результатов работы алгоритмов идентификации диктора

При обучении алгоритмов идентификации диктора реализованы 2 подхода. Первый включает в себя извлечение из речевого сигнала информативных признаков в

виде мел-частотных кепстральных коэффициентов. Далее, на основе полученных коэффициентов формировалась модель диктора с применением гауссовых смесей. Для описания модели диктора использовались 32, 64 и 128 гауссиан. В качестве метода оптимизации для переоценки параметров взят EM-алгоритм (expectation and maximization algorithm). Для инициализации моделей использовался алгоритм кластеризации k -средних. Далее, при анализе результатов будем называть данный алгоритм – RV-EM- N , где N – количество гауссиан [3–7].

Второй подход основан на применении универсальной фоновой модели (UBM, universal background model), представляющей собой модель гауссовых смесей, ранее обученную на основе записей большого числа других дикторов. Использовались реализации с 32 и 128 гауссианами. Для получения модели конкретного диктора применялся метод адаптации дикторонезависимой модели UBM. Первым шагом алгоритма является выполнение одной итерации EM-алгоритма, используя векторы признаков обучающего сигнала диктора в качестве входных данных, а параметры UBM – в качестве начальных параметров модели. После переоценки параметров выполняется их адаптация [8–10].

В табл. 2 приведены результаты работы алгоритмов распознавания диктора. Каждый метод проходил этап обучения и тестирования. Для чистоты и детальности исследования эксперимент проводился повторно четыре раза, после чего полученные данные усреднялись. Под временем работы здесь понимается продолжительность обучения и тестирования алгоритма в течение четырех циклов. Эксперимент проводился на персональной вычислительной машине с процессором AMD Phenom II X4 945 Processor 3.00 Ghz и 8 Gb оперативной памяти.

Обучающий набор данных состоял из 300 примеров – на каждый класс по 3 образца, тогда как тестовое множество состояло из 200 примеров – на каждый класс 2 образца. Результаты исследования работы алгоритмов показывают высокий уровень верности идентификации диктора, а именно более 97 %, что позволяет говорить о практической применимости подходов в реальных системах распознавания личности.

Таблица 1. Фрагмент реализованного словаря

№ строки	Число 1	Слово 1	Число 2	Слово 2	Число 3
1	483	прохождение	214	поселенка	4534
2	879	разгибатель	514	радиоточка	9147
3	532	лукавица	574	телёночек	3190
4	917	замаливание	652	риторичный	3095
5	325	артиллерия	146	разумный	1144
6	573	завербовать	119	идеология	5050
7	320	земляника	200	перепелица	7259
8	854	припевка	599	колодочный	7353
9	573	распутывать	644	кардинал	5139

Таблица 2. Результаты работы алгоритмов идентификации диктора при неоднократном повторе эксперимента ($n = 4$) с усреднением показателя верности идентификации

Алгоритм иденции	RV-EM-32	RV-EM-64	RV-EM-128	UBM-32	UBM-128
Характеристики					
Уровень верной идентификации	98,51 %	98,01 %	98,01 %	97,51 %	97,51 %
Время работы, сек	1240	2694	4942	683	2624

Для начала более детально рассмотрим результаты работы алгоритма RV-EM при использовании различного количества гауссиан – 32, 64 и 128. Видно, что применение более 32 гауссиан в рассматриваемой задаче распознавания до 100 дикторов является нецелесообразным – вычислительная сложность подхода растет, при этом точность идентификации не повышается. При анализе работы алгоритма на основе UBM также просматривается оптимальность использования 32 гауссиан.

Алгоритм идентификации диктора с применением универсальной фоновой модели UBM на основе 32 гауссиан показал точность работы в 97,51 %, уступая RV-EM-32 алгоритму 1 %. Однако метод на основе UBM-32 выигрывает у RV-EM-32 в производительности в 1,81 раза. Быстродействие объясняется тем, что используется заранее подготовленная модель, которая адаптируется для каждого диктора на этапе обучения всего за 1 шаг EM-алгоритма.

В итоге, обобщая вышесказанное, можно сделать вывод о том, что для идентификации диктора оптимальным, с точки зрения точности и требовательности к вычислительным ресурсам, является алгоритм с применением универсальной фоновой модели UBM на основе 32 гауссиан.

Исследование алгоритма распознавания Формирование базы лиц

Для тестирования алгоритма распознавания лиц сформирована база из фотографий тех же людей, которые принимали участие в записи аудиоданных. По результатам накопления сформирован набор изображений на 100 классов. Каждый класс представлял собой 3–5 изображений одного и того же человека. В итоге база насчитывала 424 цветных изображений разрешением 2448×3264 (рис. 1).



Рис. 1. Сегментированные изображения разрешением 250×250 пикселей

Особенности подготовленной базы: состоит из цветных изображений; хорошо структурирована; лица имеют различный угол поворота/наклона головы, цвет лица/волос, наличие/отсутствие очков/бороды, усов; съемка с использованием разных сцен и степени освещенности. Все эти условия приближают проведение эксперимента к реальным условиям работы системы.

Для улучшения работы системы распознавания по визуальным данным было решено сегментировать область лица, после чего привести изображения к одному разрешению 250×250 (рис. 1).

Анализ результатов работы алгоритма распознавания лица на основе сверточной нейронной сети

В качестве алгоритма распознавания лиц использовалась сверточная нейронная сеть (CNN), архитектура которой представлена на рис. 2. Сверточные нейронные сети относятся к алгоритмам глубокого обучения. Декрипторы изображений формируются за счет операции двумерной свертки, при этом сверточные фильтры формируются в процессе обучения нейронной сети. В данном исследовании использована длинная последовательность сверточных слоев, так как сверточные нейронные сети с подобной архитектурой в последнее время достигают наилучших результатов во многих задачах [11–16].

Предлагаемая сеть включает в себя 13 сверточных слоев, каждый из которых содержит линейный оператор – банк фильтров свертки (С), за которым следуют функция активации ReLU. Для дополнительного сокращения числа параметров сеть имеет 5 слоев субдискретизации (макспулинга, МП). Во всех сверточных слоях применяются фильтры размером 3×3, а в слоях субдискретизации – фильтры размером 2×2. Последние три блока являются полносвязными слоями, из которых первые два слоя имеют выходы размерностью 4096 и функцию активации ReLU. Полученный вектор передается на последний полносвязный слой, который имеет размер выхода, равный числу классов обучающей базы – 2622 и логистическую функцию активации для вычисления апостериорных вероятностей класса.

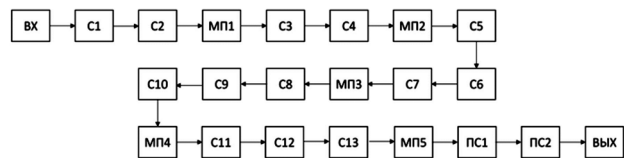


Рис. 2. Архитектура сверточной нейронной сети

Цель обучения многоклассового классификатора состоит в том, чтобы найти параметры сверточной нейронной сети, которые минимизируют значение функции потерь, при ошибке предсказания на выходе последнего слоя. Для обучения модели использовался фреймворк глубокого обучения Caffe и часть базы лиц VGG-Face [17], которая представляет собой набор изображений известных личностей по версии Internet Movie Data Base. Всего в базе содержится 2622 класса, а общее количество изображений составляет 1 891 323. Тренировочные изображения лиц перемасштабировались до размера 256×256 пикселей, а во время обучения на каждой эпохе из исходного изображения случайным образом выбирался фрагмент изображения размером 224×224 пикселей. Примеры изображений из базы VGG-Face представлены на рис. 3.

На вход сети подавались пары фотографий, которые могли комбинироваться как изображения, соответствующие одному классу, так и расходящимся классам. Так как собранная база состояла из 424 цветных изображений, то общее число всевозможных пар равнялось 89676 (комбинации из одинаковых изображений не учитыва-

лись): 702 пары изображений, имеющих совпадение класса, и 88974 пары – относящихся к разным классам. Далее, на выходе системы для каждого изображения формировался вектор признаков длиной 4096. Полученные векторы в каждой паре изображений сравнивались путем оценки евклидова расстояния d между ними, после чего результат соотносился с порогом ε . Пара изображений определялась одному и тому же классу, когда $d < \varepsilon$, и разным классам – при $d > \varepsilon$ [18, 19].



Рис. 3. Примеры изображений из обучающей базы изображений VGG dataset

Так как модель классификатора – отображение примеров в предсказанные классы, то возможны четыре варианта результатов в зависимости от объекта и предсказанного класса, которые отражены в таблице сопряженности бинарного классификатора (табл. 3).

Для оценки качества работы классификатора обычно используют ROC-кривые [20]. Данный график позволяет оценить качество бинарной классификации. Он отображает зависимость доли истинно положительных классификаций (True Positive Rate, TPR), от доли ложноположительных классификаций (False Positive Rate, FPR).

Доля истинно положительных результатов – доля найденных объектов, принадлежащих классу относительно всех объектов этого класса из тестовой выборки:

$$TPR = \frac{TP}{TP + FN}$$

Доля ложноположительных ответов показывает, сколько от общего числа реально негативных объектов оказались предсказанными неверно:

$$FPR = \frac{FP}{FP + TN}$$

В качестве основной скалярной величины, характеризующей эффективность алгоритма, выступает площадь под ROC-кривой (Area Under Curve, AUC). Она эквивалентна вероятности того, что классификатор присвоит большее значение случайно выбранному положительному объекту, чем случайно выбранному отрицательному объекту. Сама AUC оценка является агрегированной характеристикой качества классификации, не зависящей от соотношения цен ошибок. Чем больше

значение AUC, тем «лучше» модель классификации. Данный показатель часто используется для сравнительного анализа нескольких моделей. Классификатор называют случайным, если $AUC = 0,5$. Для ROC-кривой в точке (0,0) порог минимален. Идеальным случаем для классификатора является проход графика через точку (0,1) [21, 22].

На рис. 4 представлена ROC-кривая, характеризующая зависимость работы системы идентификации лица от величины ε . Порог динамически изменялся в диапазоне значений [0,085;1,220].

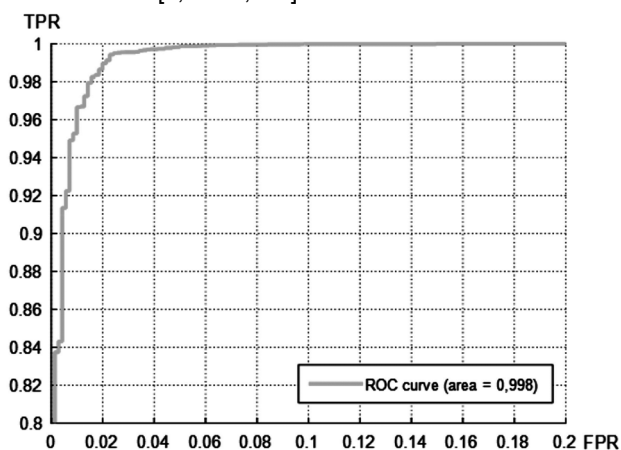


Рис. 4. ROC-кривая тестируемой сверточной нейронной сети

Из результатов видно, что площадь под ROC-кривой AUC равняется 0,998 – это свидетельствует о высокой эффективности идентификации используемой сверточной нейронной сети.

Заключение

Реализованы и исследованы алгоритмы идентификации личности по голосу и лицу. Для обучения и тестирования сформированы базы речевых сигналов и цветных изображений на 100 классов. Установлено, что использование более 32 гауссиан в представленных алгоритмах распознавания диктора является нецелесообразным. Наилучший показатель по уровню верной идентификации речевого сигнала – 98,51 %, показал алгоритм RV-EM-32. Касательно производительности, метод на основе модели UBM-32 выигрывает у RV-EM-32 в 1,81 раза. Для идентификации лица использовалась обученная ранее сверточная нейронная сеть. Ее применение на тестовой базе лиц позволило получить высокие результаты – $TPR = 0,97$ и $FPR = 0,01$. В качестве бимодальной системы распознавания личности может использоваться связка UBM-32 – CNN с уровнем верности идентификации более 97 %.

Таблица 3. Таблица несоответствий

Предсказанный класс	Фактический класс	
	Положительный (+)	Отрицательный (-)
Положительный (+)	Истинно положительный (True Positives, TP)	Ложноположительный (False Positives, FP)
Отрицательный (-)	Ложноотрицательный (False Negatives, FN)	Истинно отрицательный (True Negatives, TN)

Литература

1. Мурынин А.Б., Десятников А.А., Ковков Д.В., Лобанцов В.В., Маковкин К.А., Матвеев И.А., Чучупал В.Я. Мультимодальная биометрия – перспективное решение. Объединение алгоритмов для повышения надежности распознавания человека // Системы безопасности. 2006. № 6. С. 156–160.
2. Зализняк А.А., Гришина Е.А. Грамматический словарь русского языка. – М.: АСТ-Пресс, 2016. 800 с.
3. Первушин Е.А. Обзор основных методов распознавания дикторов // Математические структуры и моделирование. 2011. № 24. С. 41–54.
4. Садыхов Р.Х., Ракуш В.В. Модели гауссовых смесей для верификации диктора по произвольной речи // Докл. БГУИР. Минск, 2003. С. 95–103.
5. Beigi H. Fundamentals of Speaker Recognition // Springer US, Boston, 2011. 942 p.
6. Козлов А.В., Кудашев О.Ю., Матвеев Ю.Н., Пеховский Т.С. Система идентификации дикторов по голосу для конкурса NIST SRE 2013 // Труды СПИИРАН, 2013. № 2. С. 350–370.
7. Матвеев Ю.Н. Технология биометрической идентификации личности по голосу и другим модальностям // Вестник МГТУ им. Н.Э. Баумана. Сер. «Приборостроение». 2012. № 3. С. 46–61.
8. Скопинцев Я.М., Тупицин Г.С. Использование биарных масок для повышения качества закрытой текстонезависимой идентификации диктора в условиях шумов // Докл. межд. конф. «Радиоэлектронные устройства и системы для инфокоммуникационных технологий». Москва. 2014. С. 392–395.
9. Reynolds D.A., Quatieri T.F., Dunn R.B. Speaker Verification Using Adapted Gaussian Mixture Models // Digital Signal Processing. 2000. V.10, no. 1–3. P. 19–41.
10. May T., Par S., Kohlrausch A. Noise-Robust Speaker Recognition Combining Missing Data Techniques and Universal Background Modeling // IEEE Transactions on Audio, Speech, and Language Processing. 2012. Vol. 20, no. 1. P. 108–121.
11. Russakovsky O., Deng J., Su H., Krause J., Satheesh S., Ma S., Huang S., Karpathy A., Khosla A., Bernstein M., Berg A.C., Li F.F. Imagenet large scale visual recognition challenge. IJCV, 2015.
12. Goodfellow I.J., Bulatov Y., Ibarz J., Arnoud S., Shet V. Multi-digit number recognition from street view imagery using deep convolutional neural networks, 2014.
13. Simonyan K., Zisserman A. Very deep convolutional networks for large-scale image recognition // In International Conference on Learning Representations, 2015.
14. Szegedy C., Liu W., Jia Y., Sermanet P., Reed S., Anguelov D., Erhan D., Vanhoucke V., Rabinovich A. Going deeper with convolutions. CoRR, abs/1409.4842, 2014.
15. Хрящев В.В. К вопросу о выборе наилучшего лица для систем биометрической идентификации/верификации // Цифровая обработка сигналов и ее применение (DSPA-2017): докл. 19-й междунар. конф. Москва, 2017. Т. 2. С. 747–752.
16. Лебедев А.А., Хрящев В.В., Павлов В.А. Разработка алгоритма биометрической идентификации по изображению лица на основе сверточных нейронных сетей // Телевидение: передача и обработка изображений: Материалы 14-й международной конференции. Санкт-Петербург, 2017. С. 133–136.
17. Parkhi O.M., Vedaldi A., Zisserman A. Deep Face Recognition // British Machine Vision Conference. 2015.
18. Dean J., Corrado G., Monga R., Chen K., Devin M., Mao M., Ranzato M., Senior A. Large scale distributed deep networks // In Advances in Neural Information Processing Systems (NIPS). 2012. P. 1232–1240.
19. Sun Y., Chen Y., Wang X., Tang X. Deep learning face representation by joint identification-verification // In Advances in Neural Information Processing Systems. 2014. P. 1988–1996.
20. Fawcett T. An introduction to ROC analysis // Pattern Recognition Letters. 2006. Vol. 27, no. 8. P. 861–874.
21. Hernandez-Orallo J. ROC curves for regression // Pattern Recognition. 2013. Vol. 46, No. 12. P. 3395–3411.
22. Шемяков А.М., Степанова О.А., Хрящев В.В. Распознавание лиц на изображениях в условиях искажающих факторов // Цифровая обработка сигналов и ее применение (DSPA-2016): докл. 18-й междунар. конф. Москва, 2016. Т. 2. С. 983–988.

НОВЫЕ КНИГИ

Витязев С.В. Цифровые процессоры обработки сигналов / Курс лекций – М.: Изд-во «Горячая линия-Телеком», 2017 г. – 100 с.: ил.

Рассмотрены основы построения архитектур и оптимизации программного обеспечения цифровых сигнальных процессоров. Сформулированы основные задачи цифровой обработки сигналов на сигнальных процессорах. Представлено описание инструментальных и программных средств работы с цифровыми сигнальными процессорами.

Для студентов технических вузов радиотехнических и инфокоммуникационных специальностей, будет полезна преподавателям, читающим соответствующие курсы.



СРАВНИТЕЛЬНАЯ ОЦЕНКА СПОСОБОВ ОБРАБОТКИ ЦИФРОВЫХ ИЗОБРАЖЕНИЙ, ПОЛУЧАЕМЫХ С ПОМОЩЬЮ ТЕПЛОВИЗОРА ПРИ НАБЛЮДЕНИИ БЕСПИЛОТНЫХ ЛЕТАТЕЛЬНЫХ АППАРАТОВ

Зайцев А.В., к.т.н., генеральный директор ООО «Смоленский научно-инновационный центр радиоэлектронных систем «Завант», e-mail: info@zavant.ru;

Соловьев В.А., к.т.н., профессор, начальник отдела ООО «Смоленский научно-инновационный центр радиоэлектронных систем «Завант», e-mail: info@zavant.ru;

Кичулкин Д.А., начальник отдела ООО «Смоленский научно-инновационный центр радиоэлектронных систем «Завант», e-mail: info@zavant.ru;

Красавцев О.О., начальник отдела ООО «Смоленский научно-инновационный центр радиоэлектронных систем «Завант», e-mail: info@zavant.ru.

COMPARATIVE ESTIMATION OF WAYS OF PROCESSING DIGITAL IMAGES, RECEIVED BY MEANS OF THERMAL IMAGER UNDER OBSERVATION OF UNMANNED AERIAL VEHICLE

Zaitsev A. V., Solovjov V. A., Kichulkin D. A., Krasavtsev O. O.

The article considers the issue of extracting an image of a small unmanned aerial vehicle against the background of cloudy sky radiation in the infrared waves range on video frames obtained from the thermal imager output. It was proposed to perform preliminary processing of digital images to detect such air targets in order to increase the signal-to-noise ratio. The review of previously proposed methods of preliminary treatment was also made; their positive and negative aspects were singled out. Based on the results of full-scale studies, the effectiveness of each of them was evaluated. A conclusion was made about the advisability of using a combination of several methods, which will allow the operation of an optoelectronic detector of small unmanned aerial vehicles in various meteorological situations.

Key words: processing algorithm, unmanned aerial vehicle, infrared range of waves, detection, digital image, convolution, spatial differentiation, spatial filtration.

Ключевые слова: алгоритм обработки, беспилотный летательный аппарат, инфракрасное излучение, обнаружение, цифровое изображение, свертка, пространственное дифференцирование, пространственная фильтрация.

Введение

Актуальность своевременного обнаружения малоразмерных беспилотных летательных аппаратов (БЛА) в настоящее время постоянно повышается в связи с участвовавшими случаями их несанкционированного проникновения к охраняемым объектам и использования в террористических целях. Поскольку применить для обнаружения БЛА штатные радиолокационные средства не представляется возможным в связи со слабыми рассеивающими свойствами таких летательных аппаратов, целесообразно использование всесуточных оптико-электронных средств – специализированных тепловизоров.

Основными задачами тепловизионного обнаружителя малоразмерных БЛА является своевременное обнаружение, обеспечивающее требуемую дальность действия, устойчивое сопровождение цели, определение угловых координат относительно оптической оси обнаружителя, обеспечение требуемой разрешающей способности.

Задача обнаружения малоразмерного летательного аппарата с помощью всесуточного оптико-электронного

Рассматривается вопрос о выделении изображения малоразмерного беспилотного летательного аппарата на фоне излучения облачного неба в инфракрасном диапазоне волн на видеокдрах, получаемых с выхода тепловизора. Предложено для обнаружения таких воздушных целей выполнять предварительную обработку цифровых изображений с целью повышения отношения сигнал/помеха. Произведен обзор ранее предложенных способов предварительной обработки, выделены их положительные и негативные стороны. На основании результатов натурных исследований произведена оценка эффективности каждого из них. Сделан вывод о целесообразности применения комбинации нескольких способов, что позволит обеспечить работу оптико-электронного обнаружителя малоразмерных беспилотных летательных аппаратов при различных метеорологических ситуациях.

устройства – тепловизора заключается в последовательном обзоре пространства до появления в поле зрения прибора искомого объекта. С точки зрения скорости обзора желательно, чтобы тепловизор имел оптическую систему с широким полем зрения (с малым фокусным расстоянием), однако, в этом случае дальность обнаружения цели будет низкой. Это следует из анализа уравнения дальности обнаружения объектов тепловизором [1]

$$R_{\text{обн}} = \sqrt{\frac{\pi \cdot q_{\text{н}} \cdot J_{\text{ц}} \cdot \tau_{\text{об}} \cdot \tau_{\text{а}} \cdot S_{\text{р}}}{4 \cdot E_{\text{п}} \cdot m_{\text{р}} \cdot \left(\frac{f_{\text{об}}}{D_{\text{об}}}\right)^2}}, \quad (1)$$

где $J_{ц}$ – сила излучения цели, равная $B_{ц} \cdot S_{ц}$; $B_{ц}$ – энергетическая яркость цели; $S_{ц}$ – площадь проекции цели на картинную плоскость; R – расстояние до цели; $f_{об}$ – фокусное расстояние оптической системы; $D_{об}$ – диаметр оптической системы (объектива); $\tau_a, \tau_{об}$ – коэффициенты пропускания атмосферы на трассе и оптической системы; $q_{ц}$ – площадь чувствительной площадки (пикселя); m_p – отношение сигнал/шум; S_p – чувствительность по падающей мощности.

Физически это означает, что при широком поле зрения, ввиду малости размеров цели, её изображение укладывается в размеры одного пикселя (или меньше), а его яркость невелика. В такой ситуации глаз человека не способен выделить изображение цели на фоне излучения облачного неба. Для улучшения условий наблюдения целесообразно применять оптическую систему с более узким полем зрения (с большим фокусным расстоянием), тогда, при неизменной дальности, изображение цели занимает несколько пикселей и вероятность его обнаружения повышается. Вместе с тем следует учитывать, что при узком поле зрения его необходимо постоянно перемещать в процессе обзора пространства в пределах сектора ответственности, поэтому время обзора пространства увеличивается. Кроме того, с течением времени зрительный аппарат оператора утомляется, в результате чего повышается вероятность пропуска цели или уменьшение дальности её обнаружения. Для устранения этого недостатка следует автоматизировать процесс обнаружения малоразмерных воздушных целей.

В случае, близком к идеальному, когда интенсивность облачного неба в инфракрасном диапазоне невелика, например, при безоблачном небе, процесс обнаружения сводится к попикселному обходу цифрового изображения с применением операции пороговой обработки. Решение о принадлежности анализируемого элемента разрешения малоразмерной цели принимается в случае превышения яркостью пикселя некоторого порогового уровня, зависящего от интенсивности флуктуаций излучения неба и внутренних шумов приёмника лучистой энергии. Задача выбора порогового уровня для рассматриваемого случая может быть решена путём оценки отношения правдоподобия одним из способов, описанных в работах [2, 3, 4 и др.]. Вместе с тем следует учитывать, что безоблачное небо в средней полосе России не является преобладающим, часто оно наблюдается фрагментарно в сочетании с другими формами облачности. По этой причине в пределах участка оптического спектра 8-13 мкм малоразмерные БЛА необходимо обнаруживать на фоне излучения облачности, интенсивность которой порой соизмерима или превышает излучение цели. При этом облачность имеет стохастическую структуру, поэтому её излучение неравномерно распределено в пространстве, в связи с чем не представляется возможным реализовать предлагаемые в [2, 3] процедуры, т. е. получать разностное изображение, как результат вычитания фоновой составляющей из анализируемого изображения с последующей поро-

говой обработкой. Такие условия наблюдения не позволяют принимать решение о наличии цели на основе использования результатов пороговой обработки в чистом виде. Для её применения требуется подготовительный этап фильтрации цифрового изображения, направленный на повышение отношения сигнал/помеха. Сущность такой подготовительной обработки заключается в выделении изображения цели по характерным признакам, присущим БЛА, при одновременном подавлении изображения облачного неба.

Алгоритмы обработки цифровых изображений

В ряде ранее опубликованных работ [5–7] авторами предлагались алгоритмы обработки цифровых изображений, к числу которых относятся: гамма-фильтрация, свёртка цифрового изображения с пробным (эталонным) сигналом, пространственная спектральная фильтрация, селекция движущихся целей, пространственное дифференцирование изображения.

Каждый из перечисленных алгоритмов позволяет осуществить подавление излучений облачного неба, т.е. естественных помех, в разной степени.

Гамма-фильтрация цифрового изображения по строкам осуществляется в соответствии с выражением [5, 8]

$$P_j = A(F_{n,j} - F_{\min})^\gamma, \quad j = 0 \dots M, \quad (2)$$

где n – номер строки изображения; j – номер пикселя в строке; P_j – вычисляемое новое значение яркости j -го пикселя; $F_{n,j}$ – исходное значение яркости пикселя с координатами n, j ; F_{\min} – минимальное значение в анализируемом кадре; A – постоянный коэффициент (0...1); γ – параметр γ -фильтрации.

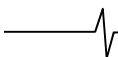
Для возможности отображения цифрового изображения на жидкокристаллическом мониторе требуется его нормирование

$$P_{ij} = \frac{250}{P_{\max}}. \quad (3)$$

При сильном сигнале цели γ -фильтрация позволяет эффективно подавлять изображения всех видов облачности, в том числе и кучевой (рис. 1). В результате обработки отношение сигнал/помеха достигает нескольких тысяч в зависимости от величины γ . В случае слабого сигнала максимум яркости в кадре приходится на облака и в результате нормировки происходит ещё большее подавление сигнала малоразмерной цели. Отношение сигнал/помеха резко снижается и, за счёт этого, дальность обнаружения БЛА оптико-электронным обнаружителем уменьшается. Этот недостаток γ -фильтрации ограничивает область её применения только случаями сильного сигнала, позволяя подчеркнуть изображение цели.

Операция свёртки направлена на выделение изображения малоразмерной воздушной цели. При этом используется известное свойство изображения такой цели, которое формируется на строке ограниченным количеством (3–6) пикселей, амплитуды яркостей в которых изменяются по закону функции Гаусса

$$P_n(j) = e^{-\left(\frac{j-j_0}{\tau}\right)^2}, \quad (4)$$



где $P_n(j)$ – яркость j -го пиксела на строке; j_0 – номер пиксела, соответствующего центру изображения цели; τ – протяжённость изображения цели вдоль строки.

Операция свёртки состоит в последовательном вычислении интеграла свёртки в цифровой форме исходного изображения с пробным сигналом (4)

$$P_{nj} = \sum_{i=j-20}^{j+20} F_{nj} e^{-\left(\frac{i-j}{\tau}\right)^2}. \quad (5)$$

Алгоритм обработки, реализующий выражение (5), позволяет подчёркивать малоразмерные участки яркости, близкие по размерам к воздушной цели. В результате отметки целей становятся на изображении более яркими, однако отношение сигнал/помеха обычно не повышается более чем на 10-15 %, зато снижается амплитуда флюктуаций яркости на строке с 10-15 до 0,9 % [5].

Недостатком данного вида обработки цифровых изображений является то, что изображения облаков (ярких протяжённых объектов) практически не подавляются (рис. 2). Изображение цели при сильном сигнале можно выделить путём дополнительной пороговой обработки

$$P_{nj} = \begin{cases} 1 & \text{при } P_j \geq P_{\text{пор}}, \\ 0 & \text{при } P_j < P_{\text{пор}}, \end{cases} \quad (6)$$

где $P_{\text{пор}}$ – уровень порога бинаризации.

Таким образом, операция свёртки оказывается целесообразной для подчёркивания изображений малоразмерных воздушных целей и снижения флюктуаций фоновых излучений. При больших дальностях уровень сигнала уменьшается, отношение сигнал/помеха снижается, поэтому дальность обнаружения цели сравнительно не велика.

Пространственная спектральная фильтрация пред-

усматривает подавление в цифровом изображении спектрального участка пространственных частот, соответствующих изображению облачных образований, и выделение области частот, формирующих изображение малоразмерной воздушной цели [6]. Исследования показали, что в горизонтальной и вертикальной плоскостях пространственные спектральные полосы, участвующие в формировании изображения малоразмерной воздушной цели, не одинаковы. По этой причине необходимо осуществлять фильтрацию изображения по строкам и столбцам разными фильтрами

$$P_r(y, f) = F_r(y, f) \cdot H_r(f), \quad (7)$$

$$P_b(x, f) = F_b(x, f) \cdot H_b(f), \quad (8)$$

где $H_r(f), H_b(f)$ – функции спектрального пространственного фильтра в горизонтальной и вертикальной плоскостях;

$F_r(y, f), F_b(x, f)$ – исходные распределения яркостей цифрового изображения в горизонтальной и вертикальной плоскостях;

$P_r(y, f), P_b(x, f)$ – результирующие распределения яркостей после фильтрации в горизонтальной и вертикальной плоскостях.

Примерами фильтрующих функций могут быть функции Баттерворта [8]

$$H_r(f) = \frac{1}{\left[1 + \left(\frac{f - f_{\text{цг}}}{f_{\text{срг}}}\right)^2\right]^{2k}},$$

$$H_b(f) = \frac{1}{\left[1 + \left(\frac{f - f_{\text{цв}}}{f_{\text{срв}}}\right)^2\right]^{2k}}, \quad (9)$$

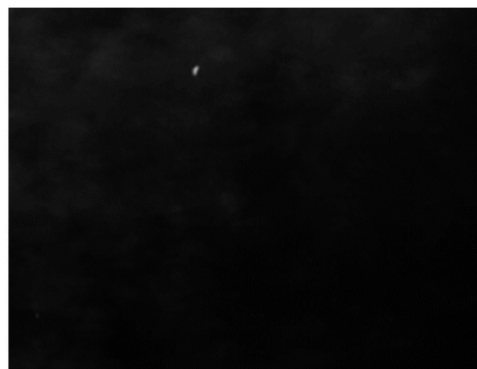


Рис. 1. Результат γ -фильтрации цифрового изображения БЛА на фоне высоко кучевой облачности ($\gamma=2,5$)



Рис. 2. Результат применения операции свёртки при обработке изображения БЛА на фоне кучевой облачности

где $f_{цг}$, $f_{цв}$ – центральные частоты полос пропускания фильтров для горизонтальной и вертикальной плоскостей; $f_{срг}$, $f_{срв}$ – частоты среза фильтров для горизонтальной и вертикальной плоскостей; k – параметр фильтра, определяющий ширину полосы пропускания.

При использовании операций (7) – (9) оказалось, что области спектральных пространственных частот, обусловленные изображением малоразмерных воздушных целей, частично пересекаются с областями, соответствующими излучению разрывов между облаками и «блестящими точками» облаков. В результате подавление кучевых облаков происходит лишь частично, в то время как другие виды облачности (слоистая, перистая, высококучевая и др.) успешно подавляются (рис. 3, а).

Неизбежным спутником спектральной пространственной обработки является некоторое размытие отметки воздушной цели, являющееся причиной снижения разрешающей способности. Устранение этого недостатка возможно путём нормировки изображения по формуле (3) и последующей пороговой обработки (бинаризации) по условию (6). Однако с увеличением дальности до цели её яркость уменьшается, и нормировка оказывается неэффективной, т. к. яркость кучевых облаков после фильтрации превышает яркость цели. По этой причине сигналы облаков превышают уровень бинаризации $P_{пор}$, поэтому наряду с изображениями целей наблюдаются и фрагменты облаков, в результате чего изображения малоразмерных целей «теряются» (рис. 3, б).

Из вышеизложенного следует, что проведение пространственной спектральной фильтрации позволяет наблюдать малоразмерные воздушные объекты на фоне

различных видов облачности кроме кучевой, наличие которой снижает дальность обнаружения целей.

Следующим направлением обработки цифровых изображений малоразмерных воздушных целей является их селекция на фоне неподвижных (малоподвижных) природных излучающих объектов. Такой подход реализуется с помощью операции черескадровой компенсации [7]

$$P_{nj}(t) = F_{nj}(t) - F_{nj}(t - \Delta t), \quad (10)$$

где Δt – интервал времени задержки вычитания.

В процессе экспериментальных исследований было установлено, что процедура (10) даёт возможность надёжного подавления излучений всех природных объектов, попадающих в поле зрения тепловизора (облаков, деревьев, кустарника, строений и т.д.). Изображение движущейся воздушной цели надёжно выделяется и обладает высокой контрастностью (рис. 4, а).

Недостатком способа черескадровой компенсации является необходимость установки тепловизора неподвижно, в противном случае при его развороте на жидкокристаллическом дисплее кроме движущихся целей наблюдаются некомпенсированные кромки облаков и контуры местных предметов (рис. 4, б). Данный способ обеспечивает самую большую дальность обнаружения малоразмерных воздушных целей.

Попытки избавиться от основного недостатка черескадровой компенсации – необходимости удержания тепловизора неподвижно – привели к необходимости более строгого учёта характерных признаков малоразмерной воздушной цели. В частности, при наблюдении цели настраке цифрового изображения наблюдается короткий сигнал вида (4). Попытка дифференцирования распреде-



Рис. 3. Результаты пространственной спектральной обработки изображений:
а – при высоко кучевой; б – при кучевой облачности



Рис. 4. Селекция движущихся воздушных целей: а – при неподвижном; б – при движущемся тепловизоре

ления яркости вдоль строки приводит к преобразованию этого сигнала в двухполярное колебание (рис. 5). Яркость вдоль строки при этом будет в пределах сигнала цели изменяться по закону

$$P_{\partial nj}(x) = \frac{\partial}{\partial x} F_{nj}(x) \approx \frac{F_{nj+1}(x) - F_{nj}(x)}{\Delta l}, \quad (11)$$

где Δl – шаг дифференцирования.

В первом приближении отличие преобразованного с помощью формулы (11) сигнала цели и облака очевидно (рис. 5). Вместе с тем анализ цифровых изображений при различных видах облачности позволил заключить, что в случае кучевой облачности с разрывами между облаками и, особенно при множественных разрывах, распределения яркости вдоль строки для цели и фрагмента облака могут оказаться сходными. По этой причине возникла необходимость уточнения принадлежности распределения яркости к малоразмерной цели. Это стало возможным благодаря сочетанию пространственного дифференцирования распределения яркости по строке изображения с последующей операцией его свёртки с ожидаемой производной распределения, обусловленного сигналом, которое можно определить путём дифференцирования выражения (4)

$$P_{\text{пл}}(j) = \frac{\partial}{\partial j} P_{\text{п}}(j) = \frac{-(2j - 2j_0) e^{-\left(\frac{j-j_0}{\tau}\right)^2}}{\tau^2}. \quad (12)$$

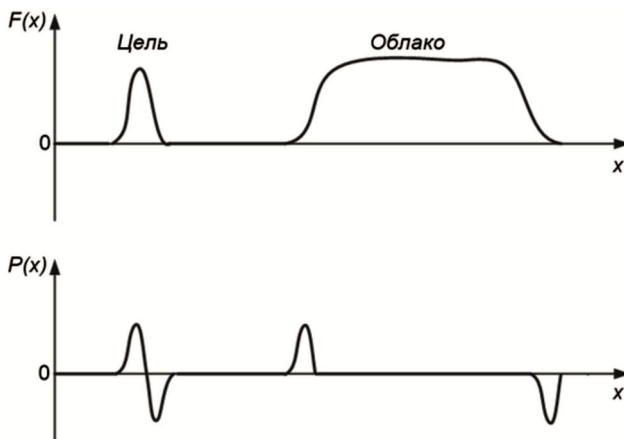


Рис. 5. Распределение яркости изображения вдоль строки изображения: $F(x)$ – до дифференцирования; $P(x)$ – после дифференцирования



Операция свёртки выражений (11) и (12) в цифровой форме осуществляется в соответствии с выражением

$$\Delta P_{nj} = \sum_{i=j-8}^{j+8} P_{\text{пл}i} \left[\frac{2P_{\text{пл}nj}(-2i - 2j_0) e^{-\left(\frac{i-(j+j_0)}{\tau}\right)^2}}{\tau^2} \right]. \quad (13)$$

Пример выделения сигнала цели с помощью процедуры (13) показан на рис. 6.

Для устранения шумовых выбросов в распределении яркости, обусловленных разрывами в облаках, необходима операция сглаживания в соответствии с выражением

$$G(j) = \frac{1}{3} \sum_{k=j}^{j+2} P_{nk}. \quad (14)$$

После устранения отрицательной части сигнала, полученного с помощью процедур (12) – (14), производится его нормировка по формуле (3) (рис. 7).

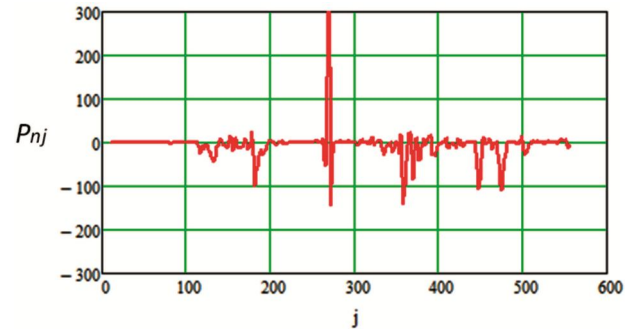


Рис. 6. Результат выделения сигнала малоразмерной воздушной цели на строке развёртки при комбинации операций пространственного дифференцирования распределения яркости и свёртки с производной ожидаемого пробного сигнала цели

В процессе наблюдения за целью возможно проведение подстройки качества изображения путём варьирования величины j_0 , определяющей положение пробного сигнала относительно сигнала цели, в пределах от 2 до 6.

Последний из рассмотренных (комбинированный) способ обработки цифровых изображений позволяет успешно подавлять сигналы всех видов облачности. При сильной изрезанности кучевой облачности возможны появления отдельных ложных отметок, обусловленных облаками, однако они успешно устраняются регулировкой величины j_0 .



Рис. 7. Результат фильтрации цифрового изображения БЛА на фоне кучевой облачности при комбинированной обработке

Оценка эффективности алгоритмов

Проведенные исследования показали, что каждый из рассмотренных выше алгоритмов не является универсальным и оказывается целесообразным при определённой фоно-целевой ситуации. В связи с этим возникает необходимость оценки эффективности каждого из алгоритмов фильтрации цифрового изображения на выходе тепловизора или их сочетания при заданных условиях. Процедура оценки включает в себя обоснование операций, которые необходимо выполнить в процессе натурных испытаний инфракрасного обнаружителя и определения числовых значений измеряемых при этом величин, для получения объективной оценки эффективности применяемого алгоритма. В свою очередь, для оценки эффективности требуется оперировать некоторыми показателями и критериями, которые также должны быть обоснованными.

Поскольку обнаружение малоразмерной воздушной цели состоит из двух последовательно выполняемых процессов (предварительной обработки изображения и принятия решения о наличии цели, т.е. собственно обнаружения) необходимо предъявить требования к каждому из них.

Для обнаружения изображения малоразмерного БЛА на фоне неба и формирования маркера с центром в точке цели (рис. 8) использован известный алгоритм сегментации движущихся объектов в видеопотоке, описанный в [9, 10]. Его сущность заключается в анализе временного распределения яркости изображения в каждом пикселе, которое подчиняется нормальному закону. Причём выделяется и анализируется несколько случайных процессов, отображаемых пикселем, что позволяет учесть временную изменчивость как характера излучения цели, так и фона облачного неба.



Рис. 8. Сопровождение обнаруженной воздушной цели

Эффективность данного алгоритма существенно зависит от отношения сигнал/шум (сигнал/помеха). Для выявления этой зависимости были проведены экспериментальные исследования, связанные с наблюдением БЛА в полёте с помощью тепловизора при различных метеорологических ситуациях. Полученные при экспериментах видеозаписи (84 записи) были использованы для анализа при различных способах обработки цифровых изображений.

В процессе сопровождения БЛА регистрировалось общее время наблюдения и суммарное время нарушений (прерываний) сопровождения, обусловленных влиянием излучения облачности. По результатам исследования определялась вероятность сопровождения, как отношение времени устойчивого сопровождения цели к общему времени наблюдения. С помощью специально разработанной программы «Statistika» (рис. 9) производилась оценка уровня полезного сигнала (сигнала воздушной цели) и уровень фонового излучения, что позволяло определять отношение сигнал/помеха.

По результатам экспериментальных исследований построен график зависимости вероятности сопровождения цели от отношения сигнал/помеха, показанный на рис. 10, из которого следует, что устойчивое сопровождение изображения малоразмерного беспилотного летательного аппарата с помощью алгоритма возможно при отношении сигнал/помеха не ниже 3-4.

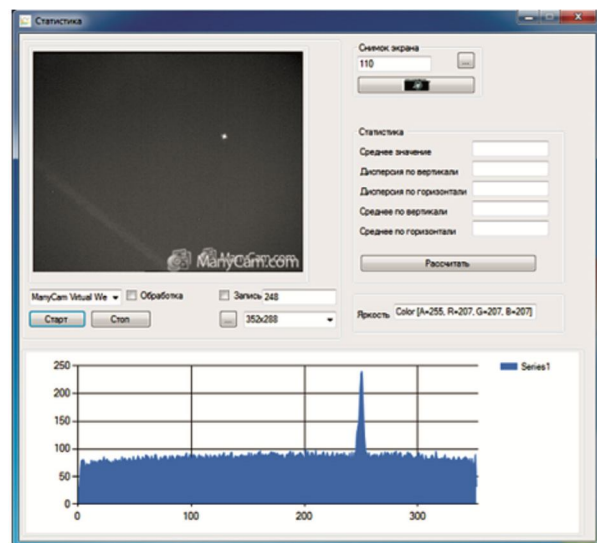


Рис. 9. Интерфейсная программа «Statistika» для оценки уровня сигнала и фона

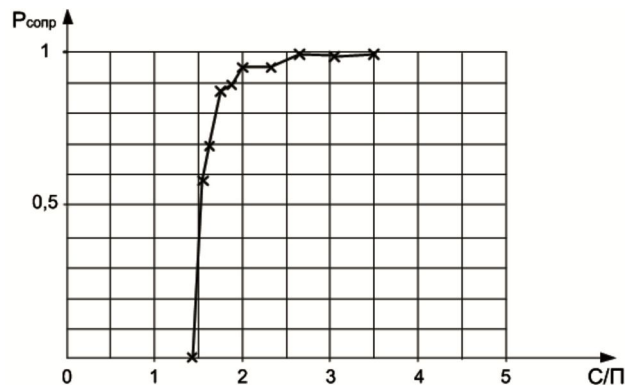


Рис. 10. Зависимость вероятности сопровождения БЛА от отношения сигнал/помеха

Таким образом, для успешного обнаружения и последующего сопровождения малоразмерных БЛА требуется с помощью операций предварительной обработки видеозаписи добиться отношения сигнал/помеха не ниже 3-4. Следовательно, одним из основных показателей эффективности цифровой обработки видеозаписей, получаемых с помощью тепловизора, является

Таблица 1. Результаты обработки цифровых изображений БЛА, получаемых с помощью тепловизора

Алгоритм обработки	Отношение сигнал/помеха					
	До обработки изображения			После обработки изображения		
	Безоблачно	Cs 1–3 балла	Cu 1–3 балла	Безоблачно	Cs 1–3 балла	Cu 1–3 балла
γ-фильтрация	1,3	1,0	0,47	1,7	1,2	0,34
Свёртка с пробным сигналом + пороговая обработка	1,78	1,15	0,67	76	76	0,34
Пространственная спектральная + пороговая обработка	1,58	1,36	0,95	250	250	1
Черескадровое вычитание при неподвижной камере	1,78	1.15	0.67	50	25	25
Черескадровое вычитание при движущейся камере	1,76	1.12	0,69	46	23	1,38
Пространственное дифференцирование + свёртка + пороговая обработка	1,71	1,17	0,87	150	37	3,3

Примечание. Cs – перисто-слоистая облачность, Cu – кучевая облачность.

отношение сигнал/помеха – N_v . Одновременно представляется возможным определить и критерий эффективности предварительной цифровой обработки – минимальный уровень отношения сигнал/помеха – $N_{v\min} = 3-4$, обеспечивающий автоматическое обнаружение и последующее сопровождение изображения БЛА.

Используя полученные при проведении экспериментов видеозаписи сопровождения БЛА, были определены численные значения отношений сигнал/помеха, достигаемые в результате проведения предварительной обработки цифровых изображений с использованием описанных выше алгоритмов на предельных дальностях на фоне безоблачного неба, перисто-слоистой и кучевой облачности 1-3 балла, которые приведены в табл. 1.

Заключение

Из результатов в таблице можно заключить, что к наиболее эффективным алгоритмам предварительной обработки цифровых изображений следует отнести черескадровое вычитание, а также комбинацию пространственного дифференцирования в совокупности с операцией свертки с производной пробного сигнала и пороговой обработкой. Последний из этих вариантов не требует в процессе наблюдения удерживать тепловизионную камеру неподвижно, что делает его более привлекательным.

Работа выполнена в рамках и на средства проекта РФФИ № 15-07-06329.

Литература

1. Соловьев В.А., Лещев М.С., Азаров В.С. Цифровая коррекция инфракрасных изображений с целью улучшения наблюдения малоразмерных летательных аппаратов // Материалы Международной научно-практической конференции «Новая наука: состояние и пути развития». Часть 2. Стерлитамак, «Агентство международных исследований», 2016. 249 с.
2. Методы компьютерной обработки изображений / Под ред. В.А. Сойфера. – М.: ФИЗМАТЛИТ, 2003. – 784 с.
3. Методы автоматического обнаружения и сопровождения объектов. Обработка изображений и управление

ние / Алпатов Б.А., Бабаян П.В., Балашов О.Е., Степашкин А.И. – М.: Радиотехника, 2008. – 176 с.

4. Алпатов Б.А., Катаев А.А. Обнаружение движущегося объекта в последовательности изображений при наличии ограничений на площадь и скорость движения объекта // Цифровая Обработка Сигналов, № 1. М.: Российское научно-техническое общество радиотехники, электроники и связи имени А.С. Попова, 2007. – С. 11–16.

5. Митрофанов Д.Г., Соловьев В.А., Зайцев А.В., Кичулкин Д.А., Азаров В.С. Выделение изображений малоразмерных воздушных объектов на тепловизионных кадрах с использованием операции свёртки // Доклады Пятой Всероссийской конференции «Радиоэлектронные средства получения, обработки и визуализации информации». М.: Российское научно-техническое общество радиотехники, электроники и связи им. А.С. Попова, 2015. С. 176-179.

6. Соловьев В.А., Зайцев А.В., Азаров В.С., Кичулкин Д.А., Красавцев О.О. Исследование числовых характеристик цифрового фильтра для подавления излучений неба при наблюдении малоразмерных воздушных целей в инфракрасном диапазоне // Доклады 18-й Международной конференции «Цифровая обработка сигналов и её применение». М., Российское научно-техническое общество радиотехники, электроники и связи им. А.С. Попова, 2016. С. 682-687.

7. Митрофанов Д.Г., Соловьев В.А., Зайцев А.В. и др. Фильтрация цифровых изображений получаемых с помощью тепловизора, при обнаружении малоразмерных воздушных объектов. Доклады Международной конференции «Радиоэлектронные устройства и системы для инфокоммуникационных технологий». М., Российское научно-техническое общество радиотехники, электроники и связи имени А. С. Попова, 2016. С. 301–306.

8. Гонсалес Р., Вудс Р. Цифровая обработка изображений. М., Издательство «Техносфера», 2006. – 1072 с.

9. Ганов В., Конушин А. Сегментация движущихся объектов в видеопотоке. Электронный ресурс: <http://cgm.computergraphics.ru/>

10. Stauffer C., Grimson W. Learning patterns of activity using real time tracking. 2000.

УДК 004.932

АЛГОРИТМ ОПРЕДЕЛЕНИЯ ДАЛЬНОСТИ ДО ОБЪЕКТА ДЛЯ СТЕРЕОСКОПИЧЕСКОЙ СИСТЕМЫ ТЕХНИЧЕСКОГО ЗРЕНИЯ РЕАЛЬНОГО ВРЕМЕНИ

*Корепанов С.Е., к.т.н., ассистент каф. Автоматики и информационных технологий в управлении (АИТУ)
Рязанского государственного радиотехнического университета (РГРТУ), e-mail: aitu@rsreu.ru;
Смирнов С.А., к.т.н., доцент каф. АИТУ РГРТУ, e-mail: aitu@rsreu.ru;
Стротов В.В., к.т.н., доцент каф. АИТУ РГРТУ, e-mail: aitu@rsreu.ru.*

ALGORITHM FOR DETERMINING THE DISTANCE TO OBJECT FOR REAL-TIME STEREOSCOPIC TECHNICAL VISION SYSTEM

Korepanov S.E., Smirnov S.A., Strotov V.V.

In this work we present the algorithm for determining the distance to object by using real-time stereoscopic technical vision system (TVS). We use the correlation based algorithms for estimating the position of the object. The choice of the algorithm for the current observation conditions is based on performance indicators. The results of experimental research performed using synthesized videos are presented.

Key words: stereoscopic system, the object coordinate estimation algorithm, SAD criterion, vision systems, performance indicators, disparity, distance.

Ключевые слова: стереоскопическая система технического зрения, алгоритм измерения координат объектов, разностная критериальная функция, системы обработки изображений, критерии работоспособности, диспаратность, дальность.

Введение

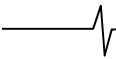
Благодаря значительному прогрессу в электронике и компьютерных системах индустрия беспилотных летательных аппаратов (БПЛА) получила значительное развитие за последнее десятилетие. БПЛА активно применяются для аэрофотосъемки, патрулирования, мониторинга различных объектов, доставки грузов, тушения пожаров и т.п.

Развитие технологий приводит к значительному удешевлению производства БПЛА и, как следствие, к увеличению их количества. Появление большого количества БПЛА, а так же тот факт, что в большинстве стран использование БПЛА практически не регламентируется, приводит к возникновению проблем, связанных с их эксплуатацией и обеспечением безопасности полетов. При этом должны быть решены такие задачи, как автономная навигация БПЛА, предотвращение столкновений, автономная посадка [1-5]. Данные задачи могут быть решены при помощи радиолокационных систем, однако данный класс аппаратуры ориентирован на крупногабаритные летательные аппараты, оснащенные мощными двигателями. В то же время, массовый сегмент современных БПЛА в основном включает в себя малоразмерные аппараты. Большинство таких БПЛА оснащены высококачественными видеодатчиками, позволяющими формировать изображения высокого контраста и с хорошим пространственным разрешением. При этом характеристики

Рассмотрен алгоритм определения дальности до объекта для стереоскопической системы технического зрения (СТЗ) реального времени. Для определения положения объекта в разных каналах СТЗ используются корреляционные алгоритмы измерения координат. Для выбора наиболее подходящего для текущих условий наблюдения алгоритма предложен критерий работоспособности. Представлены результаты экспериментальных исследований, выполненные с использованием синтезированных видеосюжетов.

современной элементной базы позволяют реализовать на борту летательного аппарата достаточно сложные алгоритмы анализа изображений [6]. В частности, использование современных ПЛИС в качестве вычислительного ядра позволяет использовать конвейерный подход к обработке поступающих изображений.

Для решения многих из перечисленных задач необходимо определять дальность до препятствий или объектов. Сделать это можно с использованием монокулярной или бинокулярных (стереоскопических) систем технического зрения (СТЗ). В случае монокулярной СТЗ определение дальности осуществляется при условии поступательного движения БПЛА за счет анализа последовательности видеок кадров, сформированной единственным видеодатчиком. При этом обязательным условием является наличие информации о скорости движения БПЛА, его координатах и ориентации в пространстве, углах поворота оптико-механической системы относительно носителя [7]. Точность определения дальности будет зависеть от датчиков углов поворота оптико-механической системы, датчиков высоты, ориентации носителя, ошибки сопоставления ключевых элементов на опорных кадрах, а также от удаленности друг от друга точек пространства, соответствующих моментам формирования опорных кадров. В стереоскопической СТЗ



используется информация от двух видеодатчиков, разнесенных между собой на некоторое базовое расстояние [5, 8]. Анализ двух изображений одного и того же объекта, позволяет определить диспаратность, т.е. величину сдвига в пикселях между изображениями объекта, формируемыми двумя камерами. Точность определения дальности зависит от характеристик камеры, базового расстояния и ошибки определения диспаратности [9].

В данной работе предложен алгоритм определения дальности до объекта для решения задачи относительной навигации двух не кооперируемых летательных аппаратов. Решение данной задачи является сложным как для БПЛА, так и для пилотируемой техники. Для определения относительного положения БПЛА относительно другого летательного аппарата необходимо сначала обнаружить данный воздушный объект, а затем осуществить слежение за ним в процессе сближения в широком диапазоне дальностей. Вопросы обнаружения и прослеживания воздушных объектов широко освещены в научной литературе, например [10-12]. В данной работе используется многоэтапный алгоритм, подробно описанный в [13].

Бортовая система управления БПЛА для предотвращения столкновения с не кооперируемым воздушным объектом должна постоянно получать информацию о дальности до объекта, т.е. задача должна решаться в режиме реального времени. Следовательно, для решения данной задачи на БПЛА должна быть установлена стереоскопическая СТЗ.

Как было отмечено выше, для определения дальности необходимо вычислить сдвиг в пикселях между изображениями объекта, полученными видеодатчиками СТЗ. Для определения диспаратности в данной работе в качестве базового алгоритма используется корреляционный алгоритм измерения координат с двумерной критериальной функцией [14]. Данный алгоритм позволяет измерять положение объекта с большой точностью, но при этом имеет высокую вычислительную сложность, и при больших размерах наблюдаемого объекта выдача информации в режиме реального времени невозможна. Ввиду того, что зачастую сближение осуществляется в хороших погодных условиях, воздушный объект будет наблюдаться на фонах с относительно равномерной яркостью. Данный факт позволяет для вычисления диспаратности использовать алгоритм, основанный на корреляционном совмещении одномерных функций. Данный алгоритм является гораздо более выгодным с точки зрения вычислений, но показывает неудовлетворительный результат при условии малого отличия средней яркости объекта от средней яркости фона. Таким образом, необходимо автоматически выбирать алгоритм вычисления диспаратности. Для автоматизации выбора будет введен критерий работоспособности для алгоритма, основанного на корреляционном совмещении одномерных функций.

Алгоритм вычисления дальности до объекта

Пусть на БПЛА установлена стереоскопическая СТЗ, включающая систему обработки изображений и два видеодатчика с идентичными параметрами. Датчики удале-

ны друг от друга на базовое расстояние, при этом их оси расположены параллельно. В поле зрения датчиков наблюдается один воздушный объект. Фон, на котором наблюдается объект, может представлять собой ясное облачное или звездное небо. Левый и правый датчики одновременно формируют два изображения $l_1(i, j, n)$ и $l_2(i, j, n)$ одинакового размера. На основании этих изображений на n -м кадре вычисляется дальность до объекта по формуле [13]:

$$D(n) = \frac{B \cdot Z}{2 \cdot Q(n) \cdot \operatorname{tg}(f/2)}, \quad (1)$$

где B – база стереосистемы в метрах, Z – разрешение изображения по горизонтали в пикселях, $Q(n)$ – значение диспаратности в пикселях, вычисленное на n -м кадре, f – ширина поля зрения камеры в градусах.

Для определения диспаратности, т.е. сдвига изображения объекта, необходимо измерить координаты объекта на изображениях с левого и правого видеодатчика. В процессе сближения с объектом изображение с левого видеодатчика используется для его обнаружения и прослеживания. На n -м кадре в качестве координат объекта можно использовать координаты, полученные от многоэтапного алгоритма [13]. Данный алгоритм, помимо координат, выдает информацию о текущих размерах объекта. Таким образом, можно сформировать эталонное изображение объекта $h_1(i, j, n)$ в левом изображении, $(i, j) \in l_1$. Поиск объекта корреляционным алгоритмом измерения координат на изображении, сформированным правым видеодатчиком, заключается в минимизации двумерной разностной критериальной функции:

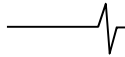
$$F_n(\alpha, \beta) = \sum_{(i, j) \in H} |l_2(i + \alpha, j + \beta, n) - h_1(i, j, n)|, \quad (2)$$

где H – множество точек эталонного изображения объекта, (α, β) – параметры, характеризующие смещение объекта вдоль осей ординат. Значение $\alpha^*(n)$ в паре $(\alpha^*(n), \beta^*(n))$, при котором достигается глобальный минимум критериальной функции (2), принимается в качестве величины, характеризующей смещение эталонного изображения объекта $h_1(i, j, n)$ относительно правого изображения

$$(\alpha^*(n), \beta^*(n)) = \arg \min_{\alpha, \beta} (F_n(\alpha, \beta)). \quad (3)$$

Значение $\beta^*(n)$ при этом может быть использовано для оценки разъюстировки видеодатчиков. Используемые в стереоскопической СТЗ видеодатчики устанавливаются таким образом, чтобы наблюдаемый объект находился на одной и той же строке на разных изображениях, т.е. при значительном отклонении $\beta^*(n)$ от нулевого значения данные измерения дальности могут быть забракованы.

Поиск эталонного изображения по всему кадру $l_2(i, j, n)$ является нецелесообразным. Обычно формируется некоторая зона поиска, координаты центра которой соответствуют центру объекта в левом изображении, а размеры выбираются из расчета предполагаемого максимального смещения объекта. При приближении к объекту диспаратность увеличивается и, как следствие,



необходимо увеличивать зону поиска, что ведет к увеличению вычислительной сложности и невозможности работы в режиме реального времени. Для того чтобы не увеличивать зону поиска, необходимо задавать положение зоны поиска с учетом диспаратности, полученной на предыдущем кадре.

Оценивание диспаратности $\alpha^*(n)$ производится с точностью до целого пикселя, однако возможно субпиксельное уточнение найденных значений. Для чего производится параболическое интерполирование разностной критериальной функции (2) в ближайшей окрестности точки минимума [15]:

$$Q(n) = x - \frac{1}{2} \times \frac{F_n(x+1, y) - F_n(x-1, y)}{F_n(x+1, y) + F_n(x-1, y) - 2 \cdot F_n(x, y)}, \quad (4)$$

где $x = \alpha^*(n)$, $y = \beta^*(n)$.

Таким образом, алгоритм вычисления дальности до объекта по поступающей паре изображений состоит из следующих шагов:

- формирование эталонного участка изображения на левом изображении;
- формирование зоны поиска эталонного участка изображения объекта интереса на правом изображении;
- вычисление целочисленной величины диспаратности с учетом диспаратности, рассчитанной на предыдущем шаге $Q(n-1)$, на основе корреляционного алгоритма измерения координат с двумерной критериальной функцией. На первом кадре $Q(n-1)$ задается нулевым значением;
- субпиксельное уточнение полученного сдвига с использованием параболического интерполирования разностной критериальной функции в окрестности минимума;
- вычисление дальности до объекта по формуле (1).

Алгоритм вычисления диспаратности, основанный на корреляционном совмещении одномерных функций

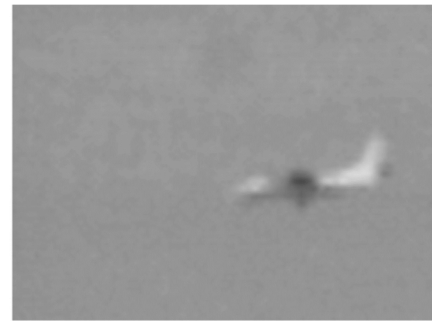
Недостатком предложенного подхода к вычислению дальности до объекта является использование корреляционного алгоритма измерения координат с двумерной критериальной функцией для определения диспаратности. Вычислительная сложность данного алгоритма растет пропорционально увеличению размера эталонного изображения объекта. При реализации алгоритма в бортовой системе обработки изображений реального времени аппаратные ресурсы ограничены количеством доступных элементов и объемом блочной памяти кристалла ПЛИС. Таким образом, при достижении размера эталона $N \times N$ пикселей выдача информации о дальности до объекта в режиме реального времени становится невозможна. В таком случае, при вычислении дальности происходит пропуск кадров и необходимо переходить к прогнозированию дальности, что ведет к снижению точности.

Относительная навигация летательных аппаратов обычно осуществляется в относительно хороших погодных условиях. Таким образом, на изображениях средняя яркость наблюдаемого воздушного объекта будет отличаться от средней яркости фона. При таких фоно-

целевых условиях для определения значения $\alpha^*(n)$ вместо двумерной критериальной функции (2) можно использовать одномерную:

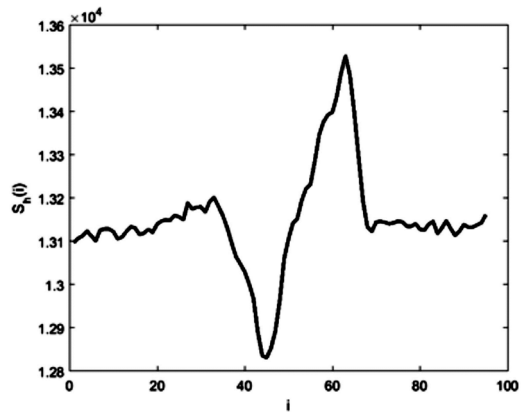


а) эталонное изображение, сформированное левым датчиком

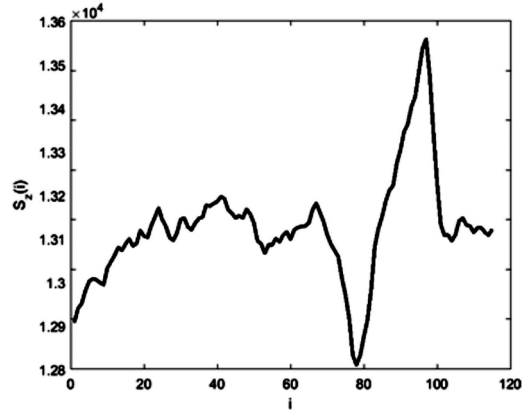


б) изображение зоны поиска, сформированное правым датчиком

Рис. 1. Изображения, формируемые датчиками стереоскопической СТЗ



а) для эталонного изображения



б) для изображения зоны поиска

Рис. 2. Результат суммирования яркостей изображений, представленных на рис. 1 по оси ординат

$$G_n(\alpha) = \sum_{i \in H} |S_z(i + \alpha, n) - S_H(i, n)|, \quad (5)$$

где $S_z(i, n) = \sum_{j \in Z} l_2(i, j)$ и $S_H(i, n) = \sum_{j \in H} h_1(i, j)$ – одно-

мерные функции, формируемые на основе суммирования изображения зоны поиска Z и эталонного изображения h_1 соответственно вдоль оси ординат. На рис. 1,а приведен пример эталонного изображения, сформированного левым видеодатчиком, а на рис. 1,б – изображения зоны поиска, сформированного правым видеодатчиком. На рис. 2,а и 2,б приведены функции $S_H(i, n)$ и $S_z(i, n)$ для этих изображений, соответственно.

Преимуществом данного алгоритма является возможность конвейеризации операции суммирования изображения, что позволяет производить данную операцию одновременно с получением изображения из видеодатчика.

Критерий работоспособности

Использование одномерной критериальной функции (5) значительно сокращает количество вычислительных операций, что позволяет производить вычисление диспаратности по формуле (3) в режиме реального времени. Однако, несоответствие текущих условий наблюдения описанным фоноцелевым условиям, приведет к неработоспособности алгоритма, основанного на корреляционном совмещении одномерных функций и, как следствие, ошибочному определению дальности. Кроме этого невозможно применение одномерной критериальной функции (5) при слежении за малоразмерным объектом.

Для вычисления диспаратности при наблюдении объекта, размеры которого не превышают $N \times N$ пикселей, целесообразно использовать корреляционный алгоритм измерения координат с двумерной критериальной функцией (2). При размерах объекта более чем $N \times N$ пикселей необходимо на каждом кадре оценивать работоспособность алгоритма, основанного на корреляционном совмещении одномерных функций, и осуществлять переключение между алгоритмами вычисления диспаратности. В случае, если оценка работоспособности алгоритма, основанного на корреляционном совмещении одномерных функций показывает его несостоятельность, необходимо продолжить использование алгоритма измерения координат с двумерной критериальной функцией, несмотря на невозможность работы в режиме реального времени. Недостающие оценки дальности должны быть получены с использованием методов экстраполяции.

Оценить эффективность алгоритма можно, введя критерий работоспособности, который будет служить индикатором качества оценки величины сдвига объекта и при этом не будет опираться на эталонные данные о местоположении объекта [16].

Эффективность применения алгоритма вычисления диспаратности, основанного на корреляционном совмещении одномерных функций, зависит от того насколько средняя яркость объекта отличается от сред-

ней яркости фона. Поэтому, в качестве критерия работоспособности можно рассматривать величину определяемую выражением:

$$KF = \begin{cases} 1, & \text{если } m_\phi + k \cdot \sigma_\phi < m_o - k \cdot \sigma_o, m_\phi < m_o \\ 1, & \text{если } m_\phi + k \cdot \sigma_\phi > m_o - k \cdot \sigma_o, m_\phi > m_o \\ 0, & \text{иначе,} \end{cases} \quad (6)$$

где m_ϕ, m_o – математическое ожидание множества точек фона и объекта, σ_ϕ, σ_o – среднеквадратичное отклонение соответствующих множеств, k – коэффициент, подбираемый экспериментальным путем (на практике используется $k = 3$). Математическое ожидание и СКО рассчитываются для левого изображения. При этом в качестве множества, содержащего точки объекта, рассматривается эталонное изображение $h_1(i, j, n)$. В качестве множества точек фона рассматривается разность между множеством эталонных точек и множеством точек, принадлежащих области, центр которой соответствует центру эталона, а размеры равны размерам зоны поиска.

В случае, когда критерий работоспособности KF принимает единичное значение, измерение диспаратности осуществляется алгоритмом, основанном на корреляционном совмещении одномерных функций.

Экспериментальные исследования

Для проведения экспериментальных исследований использовалось двенадцать синтезированных видеопоследовательностей. Сюжеты сформированы в видимом диапазоне стереоскопической СТЗ с стереобазой 25 см. Размер обрабатываемых изображений составлял 720×576 пикселей. Сюжеты содержат запись сближения БПЛА с воздушным объектом. В процессе сближения габариты наблюдаемого объекта меняются от 7×7 пикселей до размеров, при которых объект выходит за пределы поле зрения видеодатчика. Воздушный объект на сюжетах наблюдается на фоне ясного, облачного или звездного неба. Для синтеза видеопоследовательностей использовались среды моделирования Orbiter и DCS. Формирование сюжетов для экспериментальных исследований в среде моделирования позволяет получить информацию об эталонной дальности до воздушного объекта. На рис. 3 приведены примеры кадров синтезированных сюжетов.

Для оценки эффективности предлагаемого подхода оценивалась относительная ошибка определения дальности. Относительная ошибка вначале усреднялась на интервалах равных 100 метрам для каждого сюжета. После этого производилось усреднение значений ошибок на интервалах по всему набору сюжетов. В табл. 1 приведены результаты исследований определения дальности с тремя способами вычисления диспаратности использованием:

- корреляционного алгоритма измерения координат с двумерной критериальной функцией (способ № 1);
- алгоритма, основанного на корреляционном совмещении одномерных функций (способ № 2);
- переключения между способом № 1 и способом № 2 (способ № 3).



а)



б)



в)



г)

Рис. 3. Примеры кадров из тестовых сюжетов: а), б) на фоне ясного неба, в), г) на фоне ночного неба

Таблица 1. Относительная ошибка определения дальности до объекта

Дальность, в метрах	Относительная ошибка определения дальности, в %		
	Способ № 1	Способ № 2	Способ № 3
900-1000	6	16	6
800-900	9,6	13	9,6
700-800	5	8	5
600-700	6	7,5	6
500-600	5,3	5,7	5,7
400-500	4,9	5,3	5,3
300-400	3,6	4,1	4,1
200-300	2,5	2,9	2,9
100-200	1,7	1,9	1,9
50-100	1	1,5	1,5

Заключение

Результаты исследования эффективности предлагаемого алгоритма позволяют сделать вывод, что определение дальности с использованием алгоритма вычисления диспаратности, основанного на корреляционном совмещении одномерных функций, при дальностях не более 600 метров не значительно уступает в точности по сравнению с алгоритмом, в котором для вычисления диспаратности используется корреляционный алгоритм измерения координат с двумерной критериальной функцией. Использование переключения между алгоритмами вычисления диспаратности позволяет без существенной потери точности в режиме реального

времени определить дальность до объекта.

Исследования выполнены при использовании Стипендии Президента РФ СП-2598.2016.5.

Литература

1. Tirri A.E., Fasano G., Accardo D. et al. Advanced Sensing Issues for UAS Collision Avoidance // International Conference on Application and Theory of Automation in Command and Control Systems, 12-19 (2012).
2. Lai J.S., Mejias L., Ford J.J. Airborne vision based collision detection system // Journal of Field Robotics №28(2), 137-157 (2010).
3. Richard K. Smith. Seventy-Five Years of Inflight Refueling // OhneVerlagsangabe. Washington. – 1998. – 86 p.
4. Li B., Mu C., Wu B. A survey of vision based autonomous aerial refueling for Unmanned Aerial Vehicles // Intelligent Control and Information Processing (ICICIP), 2012 Third International Conference on. – IEEE, 2012. – pp. 1-6.
5. Старовойтов Е. Современные технические средства пассивной оптической локации // Современная электроника. – 2011. – №2. – С. 40-43.
6. Алпатов Б.А., Бабаян П.В., Балашов О.Е., Степашкин А.И. Методы автоматического обнаружения и сопровождения объектов. Обработка изображений и управление. – М.: Радиотехника, 2008. – 176 с.
7. Балашов О.Е. Алгоритм обнаружения наземных препятствий при маловысотных полетах // Цифровая обработка сигналов. – 2015. – №3. – С. 15-19.

8. Визильтер Ю.В., Желтов С.Ю., Бондаренко А.В., Ососков М.В., Моржин А.В. Обработка и анализ изображений в задачах машинного зрения: Курс лекций и практических занятий. – М.: Физматкнига, 2010. – 672 с.

9. Карташев В.А., Ефимов С.С. Исследование погрешностей измерения положения объекта стереосистемой с малой базой. Труды конференции «Техническое зрение в системах управления -2012»// М.: Изд. ИКИ РАН, 2012. – С. 170-172.

10. Nussberger A., Grabner H., Gool L.V. Aerial Object Tracking from an Airborne Platform // International Conference on Unmanned Aircraft Systems (ICUAS), 1284-1293, 2014.

11. Kovács L., Benedek C. Visual real-time detection, recognition and tracking of ground and airborne targets, Proceedings of SPIE-IS&T Electronic Imaging, Computational Imaging IX, vol. 7873, 1-12 (2011).

12. Balashov O., Muraviev V., Strotov V. Aerial vehicles

collision avoidance using monocular vision // Proc. SPIE 9988, Electro-Optical Remote Sensing X, 99880T

13. Alpatov B., Korepanov S., Strotov V. A composite algorithm for variable size object tracking for high performance FPGA-based on-board vision systems // SPIE Remote Sensing. – International Society for Optics and Photonics, 2014. – С. 92470A-92470A-8

14. Баклицкий В.К., Бочкарёв А.М. Методы фильтрации сигналов в корреляционно-экстремальных системах навигации. – М.: Радио и связь, 1986. – 216 с.

15. Деренченко Л.П., Троицкий В.И. Повышение точности корреляционно-экстремальных систем методом интерполяции целевой функции. // Вопросы радиоэлектроники: Сб. науч. тр. – М., 1992. – с. 16-18.

16. Корепанов С.Е., Смирнов С.А., Стротов В.В. Сравнение трех критериев работоспособности многоэталонного алгоритма оценки координат объектов // Цифровая обработка сигналов. – 2016. №3. – С.15-18.

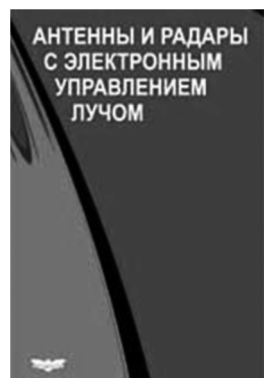
НОВЫЕ КНИГИ



Бартенев В.Г. Избранные труды (2006-2016 гг.) – М.: Изд-во «Горячая линия-Телеком», 2016 г. -478 с.: ил.

В сборник включены основные публикации за период 2006-2016 гг. Владимира Григорьевича Бартенева – доктора технических наук, профессора кафедры теоретической радиотехники и радиофизики МТУ (МИРЭА). Публикуемые труды сгруппированы по освещенной в них тематике в три раздела, представляющие основные направления научных интересов автора. Первый раздел посвящен теории и практике радиотехнических систем, во втором разделе размещено описание энергосберегающих проектов, третий раздел – это история отечественной радиоэлектроники.

Издание может быть интересно как специалистам, так и широкому кругу читателей.



Антенны и радары с электронным лучом / Под ред. А.И. Синани, Г.В. Кауфмана – М.: Изд-во «Радиотехника», 2016 г. – 234 с.: ил.

Рассмотрены наиболее значительные результаты научно-исследовательских и опытно-конструкторских работ в области антенной техники. Показана эффективность электронного управления лучом в радаров различного назначения. Проанализированы основные направления развития технологии антенных систем с электронным управлением лучом.

Для научных работников, аспирантов, инженеров-практиков, преподавателей и студентов радиотехнических специальностей.



Антипенский Р.В., Ложкин К.Ю., Поддубный В.Н., Тюлин А.Е. Аналитические сигналы в статистической радиотехнике – М.: «Радиотехника», 2016 г. – 176 с.: ил.

Рассмотрены теоретические основы и практическое применение аналитических (гильбертовских) сигналов, использование которых является развитием символического метода анализа радиотехнических процессов и цепей.

Изложены основные понятия, определения и свойства гильбертовских сигналов, даны примеры вычисления прямого и обратного преобразования Гильберта, найдены спектральные функции аналитических сигналов.

Для инженеров, научных работников, специализирующихся в области радиотехники и связи, преподавателей, аспирантов и адъюнктов.

УДК 004.932.2

АЛГОРИТМЫ ОБРАБОТКИ ИЗОБРАЖЕНИЙ ДЛЯ РЕШЕНИЯ ЗАДАЧ АНАЛИЗА ДОРОЖНОЙ ОБСТАНОВКИ

Ершов М.Д., ассистент кафедры автоматизации и информационных технологий в управлении Рязанского государственного радиотехнического университета, e-mail: aitu@rsreu.ru;

Шубин Н.Ю., к.т.н., доцент кафедры автоматизации и информационных технологий в управлении Рязанского государственного радиотехнического университета, e-mail: aitu@rsreu.ru.

IMAGE PROCESSING ALGORITHMS FOR ROAD SITUATION ANALYSIS

Ershov M.D., Shubin N.Yu.

This paper describes road situation analysis problems for traffic flow management and ensuring safety. The following image processing algorithms are proposed: road marking detection algorithm, vehicle detection and counting algorithm, stopped vehicle detection algorithm. Algorithms are designed to process images obtained from a stationary camera. The developed algorithms were implemented and tested also on the embedded platform of smart cameras. Results of experimental research of proposed algorithms are presented.

Key words: video analysis, embedded systems, image processing, object detection, background subtraction, Radon transform, multi-agent systems.

Ключевые слова: видеоаналитика, встроенные системы, обработка изображений, обнаружение объектов, оценка фона, преобразование Радона, многоагентные системы.

Введение

Появление и актуализация различных задач в области транспортной аналитики связаны с постоянно растущим числом автотранспорта на дорогах, что приводит к значительному увеличению экономических и социальных издержек. Недостаточная безопасность движения и пробки на дорогах являются прогрессирующей проблемой в городах по всему миру, и органы власти стараются свести к минимуму транспортные проблемы с помощью различных способов управления дорожным движением. Решение обозначенных проблем в условиях растущей нагрузки на транспортную инфраструктуру требует больших затрат.

Одним из возможных подходов является использование адаптивной системы управления работой светофоров на улицах города. Системы управления движением не только снижают задержки и заторы, но и решают другие проблемы:

- обнаружение происшествий;
- контроль соблюдения и фиксация нарушений ПДД;
- накопление статистической информации о транспортных потоках.

Таким образом, в настоящее время для эффективного управления дорожным движением и обеспечения безопасности всех его участников необходимо наблюдать за дорожно-транспортной обстановкой, оценивать параметры транспортных потоков, немедленно обнаруживать опасные ситуации и информировать об этом соответствующие службы. Для решения задачи анализа транспортных потоков наиболее популярными стано-

Рассматриваются задачи анализа дорожной обстановки для управления транспортными потоками и обеспечения безопасности. Предложены следующие алгоритмы обработки изображений: алгоритм обнаружения разметки, алгоритм обнаружения и подсчета автомобилей, алгоритм обнаружения остановившихся автомобилей. Алгоритмы рассчитаны на обработку изображений, получаемых стационарной камерой. Разработанные алгоритмы были реализованы и апробированы, в том числе на базе интеллектуальных камер. Представлены результаты экспериментальных исследований предложенных алгоритмов.

вятся видеодетекторы транспортных средств [1]. Они обладают следующими преимуществами:

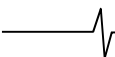
- обнаружение транспортных средств одним датчиком по нескольким полосам;
- сбор большого числа разнообразных данных о транспортном потоке;
- возможность визуального наблюдения за транспортными средствами.

Если источником информации является видеочкамера, то можно выделить два варианта реализации алгоритмов обработки данных:

На внутренней платформе камер видеонаблюдения – онлайн обработка (уменьшение количества передаваемых данных и, следовательно, сокращение требований к каналам связи и вычислительному серверу) [2].

На центральном вычислительном сервере – офлайн обработка (значительное сокращение требований к техническим характеристикам камер видеонаблюдения) [3].

Данная работа посвящена решению таких задач системы контроля дорожного движения, как обнаружение и подсчет проезжающих автомобилей, а также обнаружение остановившихся автомобилей. Кроме того, предложен алгоритм обнаружения разметки. Решение указанных задач рассматривается для неподвижной камеры, установленной над дорогой.



При разработке алгоритмов учитывалась необходимость их применения на встраиваемой платформе интеллектуальных камер видеонаблюдения. При этом уменьшается количество данных, передаваемых в систему контроля дорожного движения, однако надо помнить, что вычислительные возможности таких камер ограничены, а при обработке изображений требуются вычислительные ресурсы для учета наличия многих факторов:

- изменения освещенности сцены, вызванные тенями от облаков, деревьев, зданий;
- движущиеся тени и блики, вызванные различными положениями солнца;
- включенные фары автомобилей, блики на асфальте от фонарей уличного освещения;
- погодные условия (дождь, снег);
- различные ракурсы съемки у разных видеокамер;
- различная конфигурация наблюдаемых перекрестков и участков дорог.

Алгоритм обнаружения разметки

Для эффективной работы алгоритма обнаружения и подсчета автомобилей требуется предварительное задание зоны интереса для каждой полосы движения. С целью автоматизации этого процесса предлагается выполнять поиск разметки на наблюдаемом участке дороги.

Одним из эффективных способов обнаружения прямых линий на изображении является использование интегрального векторного преобразования Радона (ИВПР) [4], однако его применение в данной задаче ограничено следующими факторами:

- ИВПР способно обнаружить только прямые, вдоль которых лежат линии, но не концы их отрезков;
- ИВПР не способно обнаруживать кривые линии разметки;
- дисторсии, присущие некоторым видеокамерам, способны значительно искривить даже изначально прямые линии разметки.

Для преодоления перечисленных ограничений в рамках данной работы предлагается разбивать исходное изображение на небольшие фрагменты (блоки), в пределах которых кривизной линий и неточностью определения концов их отрезков можно пренебречь. Обнаруженные с помощью ИВПР линии в каждом из блоков необходимо объединить в длинные кривые или прямые линии, принимаемые за линии разметки. Для этого используется многоагентный подход [5]: каждому найденному отрезку ставится в соответствие активный агент, стремящийся к соединению с другими агентами-соседями [6]. В результате нескольких итераций взаимодействия агентов образуется устойчивая сеть потребностей-возможностей (ПВ-сеть). Её структура содержит обнаруженные линии.

Полученные, в результате работы алгоритма обнаружения разметки, предварительные оценки дорожной разметки передаются эксперту для редактирования.

Алгоритм обнаружения и подсчета автомобилей

Одной из перспективных задач является определение плотности автомобильного потока на автотрассах,

перекрестках и улицах города с целью рационального управления транспортной инфраструктурой. Системы транспортной аналитики могут использоваться как для решения задач в режиме реального времени (для адаптивного управления движением), так и для накопления статистической информации о транспортных потоках. Обработка и анализ статистической информации позволяет предложить эффективные меры для увеличения пассажиро- и грузопотока и безопасности движения: установка светофоров и регулировка их работы, изменение конфигурации перекрестка, введение дополнительных полос движения, в том числе полос для общественного транспорта, и другие.

Разработанный алгоритм обнаружения и подсчета проезжающих автомобилей предполагает задание зоны интереса на изображении – сенсора. На каждую полосу движения устанавливается свой сенсор (пример представлен в разделе «Экспериментальные исследования»). Каждый сенсор делится на две зоны (как правило, зона въезда и зона выезда), что позволяет определять направление движения проехавшего автомобиля, а также оценивать скорость движения, если были заданы размеры зон в метрах. Увеличение счетчика проехавших по полосе движения автомобилей выполняется в случае, если сенсор фиксирует следующую последовательность событий: объект обнаружен в зоне въезда, объект обнаружен в зоне выезда, объект покинул зону въезда.

Работа алгоритма начинается с оценки фона [7-9] в каждой зоне. На данной стадии важно выделить кадры, на которых отсутствует движение, для чего оценивается число «движущихся» точек на каждом кадре:

- Вычисление разности текущего и предыдущего кадра.
- Определение числа «движущихся» точек путем пороговой обработки.

Если число «движущихся» точек больше порога, зависящего от разрешения изображения, то считается, что в зоне обнаружено движение. Иначе – движения нет.

Если в течение заданного малого промежутка времени в зоне движения нет, то выбирается опорный кадр для оценки фона. Кроме того, работа зон синхронизируется с целью определения момента времени, когда автомобиль покинул зону въезда, а затем покинул зону выезда. Такая проверка необходима для исключения выбора опорного кадра во время остановки автомобиля в зоне. После нахождения опорного кадра на протяжении заданного интервала времени (как правило, несколько секунд) осуществляется проверка стабильности фона, для чего уже выполняется анализ разности опорного и текущего кадра.

Как только стадия оценки фона завершается, зона переходит в режим нормальной работы, который содержит следующие основные шаги:

1. Вычисление разности текущего кадра и фона.
2. Пороговая обработка для определения числа точек, принадлежащих к разным категориям: объект, фон, тень, засветка.

Соответственно, задаются отдельные пороги для определения точек, принадлежащих тени или засветке. Фильтрация таких точек необходима для исключения ложных обнаружений. Если же достаточно большую

площадь зоны занимают точки, отнесенные к объекту, то зона переходит в состояние «автомобиль обнаружен». Окончательное решение об обнаружении автомобиля принимается на уровне сенсора, как было описано выше.

Алгоритм обнаружения остановившихся автомобилей

Остановившееся на дороге транспортное средство может являться помехой для движения и представлять собой серьезную угрозу, в этом случае своевременное обнаружение такой ситуации позволит предупредить других участников движения и вовремя оповестить экстренные службы. Алгоритм обнаружения остановившихся автомобилей применяется ко всему изображению и основан на использовании оценки фона [7-9]. Работу алгоритма можно разделить на стадию инициализации и стадию обнаружения объектов.

Стадия инициализации заключается в первоначальной оценке фона. Для решения этой задачи применяется усредняющий фильтр: в каждом пикселе изображения определяется среднее значение яркости для заданного числа кадров [10]. Таким образом, первоначальная оценка фона осуществляется путем накопления информации о яркости пикселей в течение достаточно большого промежутка времени.

При разработке алгоритма обнаружения на основе оценки фона одной из задач является обновление этой оценки. В предлагаемом алгоритме для решения этой задачи используется экспоненциальный фильтр [11], при этом создается очередь оценок фона для разных малых промежутков времени. Имея оценку фона, обнаружение объектов можно осуществлять путем анализа модуля разности текущего изображения и фона. Нами был использован подход на основе вычитания самой ранней оценки фона из текущей оценки фона.

Таким образом, стадия обнаружения объектов включает следующие шаги:

1. Получение матрицы B_D путем бинаризации разностного изображения $D = |BG_N - BG_0|$.
2. Обработка бинарного изображения B_D .
3. Накопление информации в матрице S о продолжительности отнесения каждого пикселя к объекту.
4. Получение матрицы B_S путем бинаризации матрицы S , разметка и параметризация объектов.
5. Обновление оценок фона.

D представляет собой модуль разности текущей оценки фона BG_N и оценки фона BG_0 для самого раннего промежутка времени. Бинаризация D осуществляется путем пороговой обработки.

Следующим этапом является постобработка бинарного изображения B_D с целью связывания малых сегментов. Для этого применяются морфологические операции «открытия» и «закрытия». Также возможно использование заранее заданной маски для обнуления в матрице B_D значений пикселей на участках, не принадлежащих дорожному полотну.

Для учета временного порога обнаружения создается матрица S , хранящая для каждого пикселя число

кадров, в которых пиксель был классифицирован как принадлежащий остановившемуся объекту. Обновление матрицы S выполняется на каждом кадре на основе информации, содержащейся в бинарном изображении B_D .

Бинарное изображение B_S является результатом пороговой обработки матрицы S . Порог выставляется соответственно требуемому времени обнаружения. Наконец, выполняется процедура разметки и параметризации выделенных объектов с использованием матрицы B_S . Данная процедура обеспечивает присвоение каждому объекту уникального номера и вычисление таких параметров, как координаты и площадь. Объекты слишком малого размера отбрасываются.

Для обновления текущей оценки фона BG_N на вход экспоненциального фильтра подается текущее изображение I_N . В результате BG_N будет накапливать информацию об остановившихся объектах.

Обновление ранних оценок фона $BG_i (i = 0..N-1)$ осуществляется на основе маски. В качестве маски выступает бинарное изображение B_D . Для пикселей, принадлежащих объектам, значение яркости усредняется по всем оценкам фона $BG_i (i = 0..N-1)$. Через заданный малый промежуток времени очередь оценок фона сдвигается, самая ранняя оценка отбрасывается ($BG_0 = BG_1, \dots, BG_{N-1} = BG_N$).

Экспериментальные исследования

Экспериментальные исследования разработанных алгоритмов и программного обеспечения проводились как с использованием ранее записанных натуральных видеосюжетов, так и на внутренней платформе видеокamera при наблюдении за разными участками дорог в режиме реального времени. Съемки выполнялись при разных погодных условиях и в разное время суток. Длина видеопоследовательностей составляет от 10 минут до 8 часов. Размер обрабатываемых изображений от 360×288 до 1280×720 пикселей. Работа алгоритмов проверялась на таких сетевых камерах, как AXIS P1354 и P1365, а также на видеокодерах AXIS Q7424-R.

Исследования алгоритма обнаружения разметки проводились на кадрах, взятых из 11 видеосюжетов, содержание которых отличается как местом, так и условиями наблюдения (включая погодные).

На рис. 1 приведены примеры работы алгоритма обнаружения разметки. В табл. 1 представлены результаты экспериментальных исследований.

Таким образом, средняя эффективность обнаружения линий разметки составляет около 79 %.

На рис. 2 представлены примеры работы алгоритма обнаружения и подсчета автомобилей: изображения получены с разных камер при наблюдении за участками дорог с 3 и 5 полосами движения соответственно.

На рис. 3 представлены примеры работы алгоритма обнаружения остановившихся автомобилей. Во-первых, в солнечный день (продолжительность остановки – более 1 минуты), на протяжении данной видеопоследовательности изменялись условия освещения из-за присутствия облаков. Во-вторых, в ночное время при освещении фонарей (продолжительность остановки – более 10 секунд).

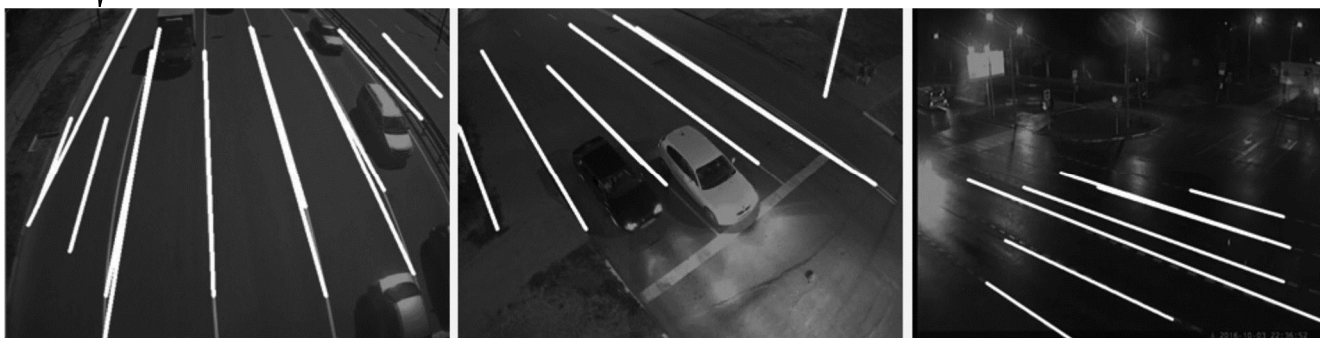


Рис. 1. Примеры автоматического обнаружения разметки

Таблица 1. Результаты исследований алгоритма обнаружения линий разметки

№ сюжета	1	2	3	4	5	6	7	8	9	10	11
Число верных обнаружений	3	5	6	5	6	5	4	5	7	4	4
Число пропусков	2	0	2	1	0	2	0	1	5	2	2

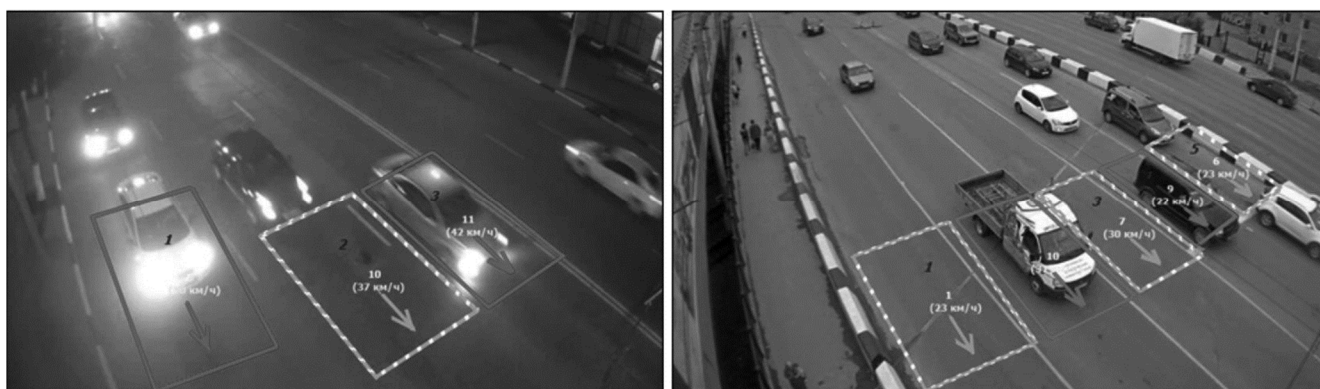


Рис. 2. Примеры обнаружения и подсчета автомобилей



Рис. 3. Примеры обнаружения остановившихся автомобилей

В табл. 2 представлены обобщенные результаты экспериментальных исследований алгоритмов обнаружения автомобилей (ОА) и обнаружения остановившихся автомобилей (ООА). Используются следующие обозначения характеристик эффективности работы:

1. P_o – доля верно обнаруженных объектов.
2. P_I – доля ошибок первого рода (ложное обнаружение).
3. P_{II} – доля ошибок второго рода (пропуск объекта).

Таблица 2. Результаты исследований алгоритмов обнаружения автомобилей

Алгоритм	Характеристики		
	$P_o, \%$	$P_I, \%$	$P_{II}, \%$
ОА	97	2	3
ООА	100	7,5	0

Заключение

В данной работе рассмотрены актуальные задачи анализа дорожной обстановки. Были представлены алгоритмы обработки видеопоследовательностей, получаемых оптическим датчиком при наблюдении за перекрестком или иным участком дороги. Данные алгоритмы позволяют выполнять поиск разметки, обнаруживать проезжающие по полосам движения автомобили, оценивать некоторые параметры транспортных потоков, обнаруживать и локализовать остановившиеся транспортные средства. Предложенные алгоритмы также реализованы и апробированы на базе интеллектуальных камер.

Эффективность алгоритмов и разработанного программного обеспечения экспериментально подтверждена. После анализа результатов исследований мы можем

заклЮчить, что представленный программный комплекс позволяет решать поставленные задачи в режиме реального времени при различных условиях наблюдения.

Работа выполнена при поддержке гранта Президента РФ для государственной поддержки молодых российских учёных МК-5331.2016.9.

Литература

1. Al-Smadi M., Abdulrahim K., Salam R.A. Traffic Surveillance: A Review of Vision Based Vehicle Detection, Recognition and Tracking // International Journal of Applied Engineering Research. – 2016. – Vol. 11 (1). – P. 713-726.
2. Alpatov B., Babayan P., Ershov M., Strotov V. The implementation of contour-based object orientation estimation algorithm in FPGA-based on-board vision system // Proc. SPIE, High-Performance Computing in Geoscience and Remote Sensing VI. – SPIE, 2016. – Vol. 10007. – 8 p.
3. Ершов М.Д. Применение алгоритмов обработки изображений для оценки параметров транспортных потоков // Новые информационные технологии в научных исследованиях: материалы XXI Всеросс. науч.-техн. конф. – Рязань: РГРТУ, 2016. – С. 259-261.
4. Babayan P., Shubin N. Line detection in a noisy environment with weighted Radon transform // Proc. SPIE 9024, Image Processing: Machine Vision Applications VII, 902409. – 2014. – 6 p.

5. Vittikh V.A., Skobelev P.O. The multi-agent models of interaction in demand-resource networks // Automatica and Telemechanica 1. – 2003. – P. 177-185.

6. Babayan P., Shubin N. Detection of curved lines and estimation of their parameters on images // 6th Mediterranean Conference on Embedded Computing (MECO), Bar. – 2017. – P. 1-4.

7. Алпатов Б.А. Оптимальное оценивание параметров движущегося объекта в последовательности изображений // Автометрия. – 1994. – № 2. – С. 32-37.

8. Алпатов Б.А., Бабаян П.В. Разработка алгоритма слежения за фоновым изображением для видеокomпьютерной системы обнаружения и определения координат движущихся объектов // Проблемы математического моделирования и обработки информации в научных исследованиях: Сборник научных трудов. – Рязань: РГРТА, 2003. – С. 3-15.

9. Benezeth Y., Emile B., Laurent H., Rosenberger C. Review and evaluation of commonly-implemented background subtraction algorithms // International Conference on Pattern Recognition. – 2008. – 4 p.

10. Piccardi M. Background subtraction techniques: a review // Proc. of IEEE International Conference on Systems, Man and Cybernetics. – 2004. – P. 3099-3104.

11. Heikkila J., Silven O. A real-time system for monitoring of cyclists and pedestrians // 2nd IEEE Workshop on Visual Surveillance. – 1999. – P. 74-81.

НОВЫЕ КНИГИ



Литвин С.М., Попов О.Б., Чернышева Т.В. *Аудиопроекторная обработка сигналов звукового вещания* – М.: Изд-во «Горячая линия-Телеком», 2017 г. – 222 с.: ил.

Звуковой сигнал, передаваемый по каналам звукового вещания, был и будет востребован в общем ряду с большим количеством альтернативных источников информации. Рассмотрены основные характеристики сигналов и каналов звукового вещания, базовые процедуры, используемые при обработке звукового вещательного сигнала (ЗВС); основные алгоритмы обработки сигнала в канале передачи. Проведен анализ и оценка искажений сигнала на всех этапах его передачи – от первичного цифрового преобразования до аудиопроекторной обработки. Особое внимание уделено алгоритмам обработки и проблеме объективной оценки качества ЗВС в системах, не нормируемых в рамках современного метрологического обеспечения.

Для студентов, обучающихся по направлениям подготовки бакалавров и магистров 11.03.02 и 11.04.02 – «Инфокоммуникационные технологии и системы связи», 11.03.01 и 11.04.01 – «Радиотехника», 10.05.02 – «Информационная безопасность телекоммуникационных систем».



Горшков Ю.Г. *Обработка речевых и акустических биомедицинских сигналов на основе вейвлетов*. – М.: Изд-во «Радиотехника», 2017 г. – 240 с.: ил.

Изложены новые перспективные методы речевых и акустических биомедицинских сигналов на основе вейвлетов, аппаратно-программные средства для решения прикладных задач криминалистического исследования фонограмм, оценки эмоционального состояния человека, шумочистки и засекречивания речи, а также построения биомедицинских систем диагностики заболеваний сердца и легких. Основное внимание уделено разработке вейвлет-технологий высокоточного анализа речевых и акустических биомедицинских сигналов. Впервые предложены решения многоуровневого вейвлет-преобразования, обеспечивающие визуализацию нестационарных сигналов с повышенным частотно-временным разрешением. Представлена система телемедицины нового поколения с засекречиванием акустических биомедицинских сигналов.

Для ученых и инженеров, работающих над созданием средств обработки биомедицинских сигналов на новых принципах. Может быть полезна преподавателям, аспирантам и студентам технических вузов.

СИНТЕЗ ПОЛНОСТЬЮ СФОКУСИРОВАННЫХ ИЗОБРАЖЕНИЙ ПРИ НАЛИЧИИ ШУМОВ

Носков А.А., аспирант кафедры инфокоммуникаций и радиофизики Ярославского государственного университета им. П. Г. Демидова, e-mail: noskoff.andrey@gmail.com;

Волохов В.А., к.т.н., доцент кафедры инфокоммуникаций и радиофизики Ярославского государственного университета им. П. Г. Демидова, e-mail: volokhov@piclab.ru;

Приоров А.Л., д.т.н., доцент кафедры инфокоммуникаций и радиофизики Ярославского государственного университета им. П. Г. Демидова, e-mail: andcat@yandex.ru;

Апальков И.В., к.т.н., доцент кафедры инфокоммуникаций и радиофизики Ярославского государственного университета им. П. Г. Демидова, e-mail: ilya@apalkoff.ru.

FULL-FOCUSED IMAGE FUSION IN THE CONDITIONS OF NOISE

Noskov A.A., Volokhov V.A., Priorov A.L., Apalkov I.V.

The implementation and research of the algorithm for full-focused image fusion in the presence of noise are presented. Three variants of images with noise fusion are considered: without using of preprocessing and postprocessing, with using of prefiltering of source images and with using of postfiltering of the fused image. The database of test scenes which created by the authors was used for full-focused image fusion algorithm tests. Additive white Gaussian noise was considered as a noise model, and a two-stage digital image processing scheme, based on principal components analysis was used as a filtering algorithm. Quantitative and visual results which are demonstrate the main features of the proposed algorithm are shown.

Key words: image filtering, machine learning, principal components analysis, full-focused image, Laplacian pyramid, Gaussian pyramid, cellular automata, image fusion, quality assessment.

Ключевые слова: фильтрация изображений, машинное обучение, анализ главных компонент, полностью сфокусированное изображение, пирамида Лапласа, пирамида Гаусса, клеточный автомат, смешивание изображений, оценка качества.

Введение

Переход от аналоговой к цифровой фотографии является большим шагом вперед. Цифровая фотография [1] открыла много новых возможностей, к которым можно отнести мгновенный предварительный просмотр снятых изображений, их быстрое редактирование, возможность простой записи видеопоследовательностей и т.д. На данный момент мегапиксельное разрешение цифровых фотокамер позволяет создавать изображения высокого качества для широкого круга потребителей и профессиональных приложений. Некоторые исследователи считают, что следующим большим шагом вперед в области формирования и обработки визуальной информации является вычислительная фотография [2], которая расширяет границы традиционной цифровой фотографии. Это обеспечивается за счет возможности записи намного большего объема информации о захватываемой сцене, а также более качественной обработки этой информации впоследствии.

Одним из важных направлений, рассматриваемых в области вычислительной фотографии, является проблема синтеза полностью сфокусированных изображений [3-7]. Суть задачи заключается в созда-

Представлены реализация и результаты исследования алгоритма синтеза полностью сфокусированных изображений при наличии шумов. Рассмотрены три варианта объединения зашумленных изображений: без предобработки и постобработки, с использованием предфильтрации объединяемых изображений, с использованием постфильтрации объединенного изображения. Для тестирования предложенного алгоритма построения полностью сфокусированных изображений использовалась созданная авторами база тестовых изображений сцен. В качестве модели шума принят аддитивный белый гауссовский шум, а в качестве алгоритма фильтрации – двухэтапная схема обработки цифровых изображений на основе анализа главных компонент. Приведены численные и визуальные результаты, демонстрирующие основные особенности предложенного алгоритма.

нии алгоритма, позволяющего выполнить объединение нескольких изображений фиксированной сцены, обладающих ограниченной глубиной резкости и сформированных с различным фокусным расстоянием. Каждое из таких изображений имеет регионы, которые можно характеризовать как размытые и, как следствие, не пригодные для дальнейшей обработки системой технического зрения или человеком, а также области, находящиеся в фокусе, которые содержат полезную информацию об окружающей сцене. Полностью сфокусированным называется изображение, содержащее в себе все находящиеся в фокусе регионы исходных изображений, пригодные для дальнейшей обработки. Таким образом, задача формирования полностью сфокусированных изображений сводится к поиску и объединению этих регионов. Необходимо отметить, что, как правило, в научно-

технической литературе данная задача рассматривается в отрыве от проблемы возникновения шумов, появляющихся в процессе формирования цифрового изображения и оказывающих негативное влияние на процедуру построения полностью сфокусированных изображений. Поэтому в настоящей работе рассматривается подход, позволяющий выполнить построение полностью сфокусированных изображений при наличии шумового воздействия. Суть подхода заключается в использовании алгоритмов префильтрации (фильтрация до объединения зашумленных изображений) или постфильтрации (фильтрация после объединения зашумленных изображений) с целью формирования качественного полностью сфокусированного изображения.

Фильтрация изображений

На сегодняшний день алгоритмы фильтрации цифровых изображений [8] широко применяются в области современной науки и техники и имеют множество практических приложений. Достаточно интересным подходом к задаче фильтрации изображений является использование методов машинного обучения [9, 10]. В рассматриваемой работе при решении задачи шумоподавления использовалась двухэтапная схема фильтрации изображений на основе анализа главных компонент (АГК) [11, 12]. Ее краткое описание представлено ниже в предположении, что анализируемое цифровое изображение x искажено аддитивным белым гауссовским шумом (АБГШ) n с нулевым математическим ожиданием и среднеквадратическим отклонением σ .

1. Первый этап обработки

1. Пусть среднеквадратическое отклонение шума σ входного зашумленного изображения $y = x + n$ является известным.

2. Разбить входное зашумленное изображение на совокупность перекрывающихся блоков. Внутри каждого из них можно выделить: *область обучения, область фильтрации и область наложения блоков*. Размеры рассматриваемых областей могут варьироваться.

3. Внутри области обучения выбрать всевозможные квадратные блоки (*векторы обучения*). Последние, будучи представленными в виде векторов-столбцов, позволяют сформировать некоторую выборочную матрицу, столбцами которой являются рассматриваемые векторы-столбцы.

4. На основе выборочной матрицы, предварительно центрировав ее, составить ковариационную матрицу, для которой найти собственные числа и соответствующие им собственные векторы (*главные компоненты* данных, заключенных в центрированной выборочной матрице). Итогом данного этапа является построение матрицы ортогонального преобразования.

5. Найти проекции (коэффициенты преобразования) множества векторов, заключенных в центрированной выборочной матрице, на множество собственных векторов, найденных на шаге 4.

6. Осуществить обработку полученного множества проекций с использованием линейной среднеквадратической оценки [10].

7. На основе множества обработанных данных восстановить оценку выборочной матрицы незашумленных данных, а на основе последней – отдельную обработанную область на изображении. При этом, во-первых, восстанавливается область обучения путем помещения в нее с учетом наложения векторов обучения в соответствующие им пространственные позиции. Векторы обучения, находясь внутри оценки выборочной матрицы незашумленных данных в виде векторов-столбцов, снова преобразовываются в квадратные блоки перед вложением в область обучения. Необходимо отметить, что область наложения векторов обучения усредняется с использованием арифметического усреднения. Во-вторых, после восстановления области обучения из нее выделяется область фильтрации меньших размеров. Повторяя аналогичную операцию для остальных областей фильтрации, с учетом их наложения, можно обработать зашумленное изображение целиком и получить первичную «грубую» оценку \hat{x}^I неискаженного изображения x . При этом обработанные области фильтрации вкладываются в соответствующие пространственные позиции изображения \hat{x}^I , а их область наложения усредняется арифметически.

II. Второй этап обработки

1. Используя зашумленное изображение y , повторить шаги 2-5, рассмотренные на первом этапе обработки. При этом устанавливаются другие размеры областей обучения, областей фильтрации и областей наложения, а также векторов обучения.

2. Осуществить обработку полученного множества проекций с использованием эмпирического фильтра Винера, представленного в области главных компонент и вычисленного на основе первичной оценки \hat{x}^I .

3. Повторяя рассуждения, изложенные на шаге 7 первого этапа обработки, получить вторую, «точную» оценку \hat{x}^{II} неискаженного изображения x .

На рис. 1 представлена блок-схема описанного алгоритма фильтрации цифровых изображений. Необходимо отметить, что при обработке цветных RGB-изображений данная схема фильтрации применяется поканально для каждой цветовой компоненты.

На рис. 2 представлен пример фильтрации цифровых изображений из рассматриваемой авторами настоящей работы базы данных, которая кратко описана ниже. Каждое из изображений сцены «Робот», сформированное с разным фокусным расстоянием, зашумлено АБГШ со среднеквадратическим отклонением $\sigma = 35$. Дополнительно на рис. 2 указаны численные оценки ПОСШ, дБ (пиковое отношение сигнала к шуму) [13] и КСП (коэффициент структурного подобия) [14] для зашумленных и восстановленных изображений.

Построение полностью сфокусированных изображений

Выше было отмечено, что синтез полностью сфокусированных изображений представляет собой процесс получения одного изображения из нескольких исходных.

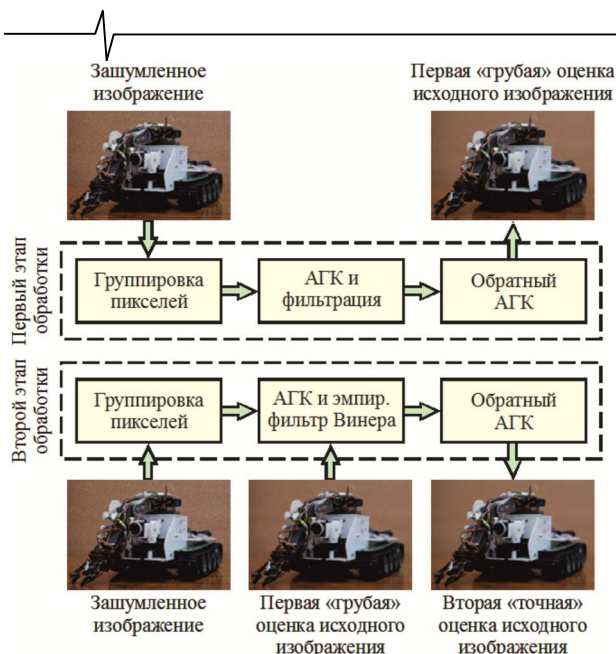


Рис. 1. Блок-схема обработки цифрового изображения с использованием двухэтапной схемы фильтрации изображений на основе анализа главных компонент [11, 12]

Получаемое изображение несет в себе больше информации о фотографируемой сцене, чем каждое из исходных по отдельности. Такое изображение может быть более удобным для дальнейшей работы с ним человеком или автоматической системой обработки изображений. В настоящей работе при решении задачи построения полностью сфокусированных изображений использовался алгоритм на основе клеточных автоматов [15, 16] и пирамид изображений [17], представленный в [7]. В предположении, что обрабатываемые алгоритмом данные заключены в нескольких изображениях фиксированной сцены, обладаю-

щих ограниченной глубиной резкости и сформированных с различным фокусным расстоянием, ниже кратко описаны основные шаги данного алгоритма. Блок-схема описываемого алгоритма представлена на рис. 3.

I. Первый этап обработки

1. Вычислить метрику сфокусированности (МС) для каждого пикселя на каждом из N исходных изображений x_k фиксированной сцены и сформировать N матриц, содержащих вычисленную метрику. Индекс k здесь является номером изображения, который изменяется от 1 до N .

2. Рассчитать максимальное значение метрики сфокусированности для каждой из N матриц, полученных на шаге 1.

3. Выполнить пороговую бинаризацию матриц, рассчитанных на шаге 1, с порогом, вычисленным на основе найденных на шаге 2 максимальных значений метрики сфокусированности и значения α , являющегося параметром алгоритма.

4. Создать матрицу меток, содержащую информацию о том, какая пара координат (i, j) пикселей из N изображений фиксированной сцены должна участвовать в формировании итогового полностью сфокусированного изображения. Формирование матрицы меток осуществляется с использованием данных на выходе пороговой бинаризации (шаг 3).

5. Произвести коррекцию матрицы меток с использованием клеточного автомата.

II. Второй этап обработки

1. Сформировать бинарные маски m_k на основе матрицы меток, полученной на шаге 5 первого этапа обработки. Количество бинарных масок равно N , а их размерность совпадает с размерностью исходных изображений фиксированной сцены. Бинарные маски необ-

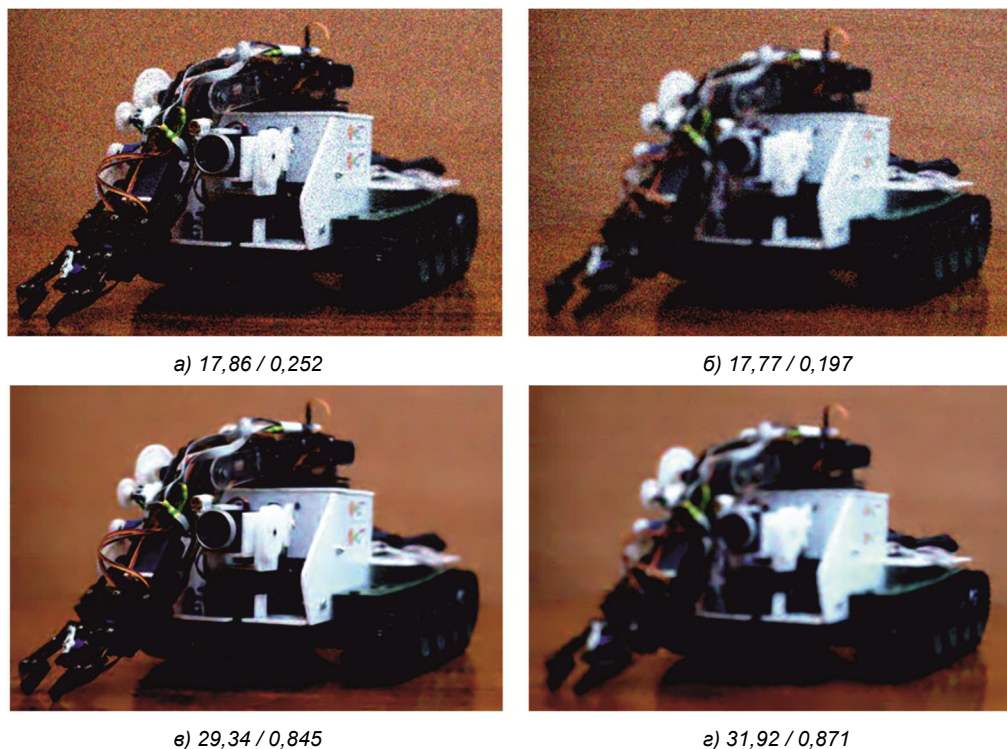


Рис. 2. Примеры изображений для сцены «Робот»: а) и б) зашумленные изображения ($\sigma = 35$); в) и г) соответствующие им восстановленные изображения

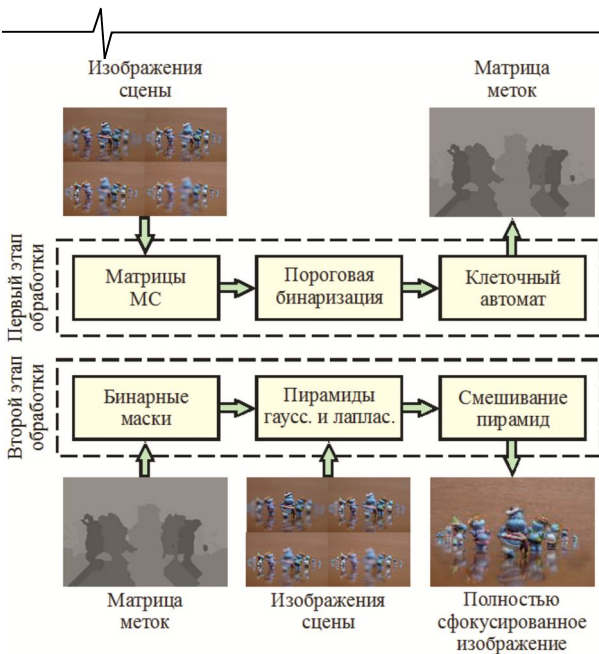


Рис. 3. Блок-схема алгоритма синтеза

полностью сфокусированных изображений на основе клеточных автоматов и пирамид изображений [7]

ходимы для реализации последующих шагов объединения исходных изображений фиксированной сцены в одно полностью сфокусированное, которое выполняется с использованием пирамид гауссианов и лапласианов [17].

2. Дополнить исходные изображения фиксированной сцены и соответствующие бинарные маски до длины ближайшей степени двойки по строкам и по столбцам. Данная процедура необходима для выполнения целочисленного деления на 2 при децимации изображений в ходе построения пирамид гауссианов и лапласианов.

3. Создать пирамиду лапласианов LP_k для каждого исходного изображения фиксированной сцены, задавая необходимое число уровней разложения.

Для цветных RGB-изображений формирование пирамиды лапласианов выполняется поканально.

4. Итерационно выполнить смешивание пирамид лапласианов LP_k с целью формирования гибридной пирамиды лапласианов GLP , обратное преобразование от которой позволит сформировать полностью сфокусированное изображение. Для проведения данной процедуры необходимо выполнить следующие шаги:

4.1. Ввести в рассмотрение интегральную маску im , обладающую размерностью бинарных масок m_k . Выполнить инициализацию начальных значений интегральной маски и гибридной пирамиды лапласианов значениями m_1 и LP_1 соответственно, вычисленными для первого изображения фиксированной сцены.

4.2. Создать пирамиду гауссианов GP^{i-1} для текущей интегральной маски im^{i-1} . Здесь индекс i описывает номер итерации, на которой происходит формирование гибридной пирамиды лапласианов.

4.3. Создать гибридную пирамиду лапласианов по следующему правилу:

$$GLP^i = GLP^{i-1} \cdot GP^{i-1} + LP_i \cdot (1 - GP^{i-1}).$$

4.4. Обновить текущую интегральную маску с использованием следующего выражения:

$$im^i = im^{i-1} + m_i.$$

4.5. Повторить шаги 4.2-4.4 до тех пор, пока не произойдет объединение пирамид лапласианов LP_k для всех исходных изображений фиксированной сцены. Получить гибридную пирамиду лапласианов.

5. Создать итоговое полностью сфокусированное изображение путем обратного преобразования гибридной пирамиды лапласианов.

На рис. 4 представлены примеры полностью сфокусированных изображений, синтезированных с использованием алгоритма [7] для групп изображений четырех сцен, рассматриваемых в работе.

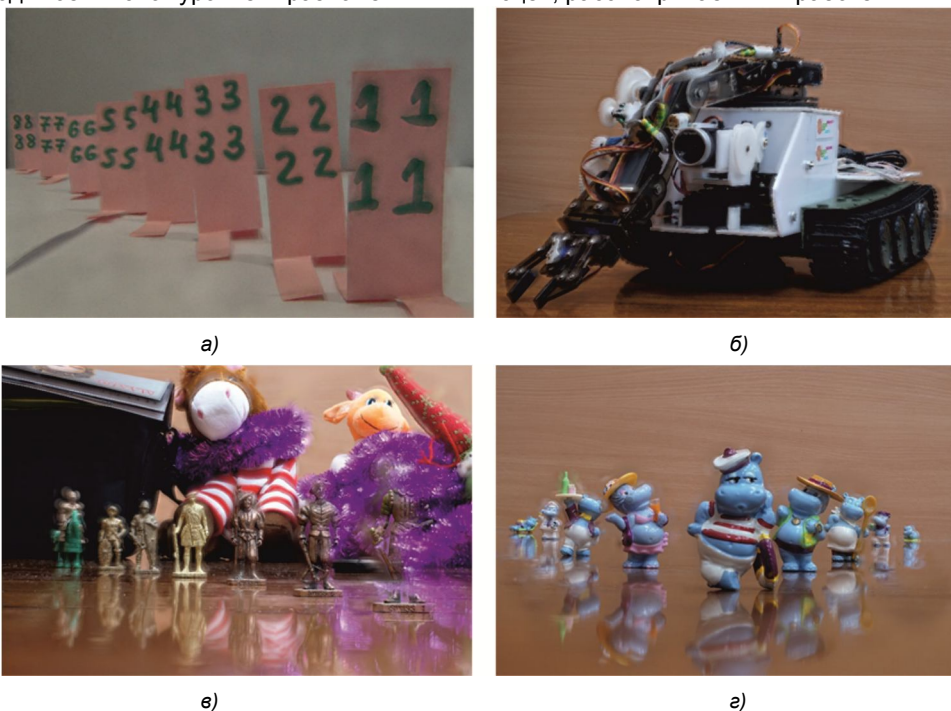


Рис. 4. Примеры сформированных полностью сфокусированных изображений для различных тестовых сцен:

а) «Числовые карточки»; б) «Робот»; в) «Солдаты»; г) «Игрушки»

Результаты моделирования

При выполнении процедуры моделирования использовались четыре группы цветных RGB-изображений фиксированных сцен:

1. «Числовые карточки», 7 изображений, разрешение 892×592 .
2. «Робот», 8 изображений, разрешение 786×523 .
3. «Солдаты», 8 изображений, разрешение 786×523 .
4. «Игрушки», 5 изображений, разрешение 786×523 .

Тестовые изображения получены в ходе выполнения настоящей работы при использовании закрепленного на штативе цифрового фотоаппарата, имеющего возможность изменения фокусного расстояния. Сформированные группы изображений позволили с использованием описанного выше алгоритма [7] синтезировать полностью сфокусированные изображения, представленные на рис. 4, которые взяты за эталонные при проведении исследования качества работы алгоритма [7] в присутствии шумов. В табл. 1 приведены численные значения ПОСШ, дБ/КСП для полностью сфокусированных изображений при различных значениях σ и различных подходах к фильтрации изображений.

Для проведения исследования рассмотрены три подхода к объединению изображений.

Подход 1. Исходные изображения сцен зашумлялись АБГШ с фиксированным σ и затем объединялись, используя алгоритм [7].

Подход 2. Исходные изображения сцен зашумлялись АБГШ с фиксированным σ , подвергались фильтрации, используя двухэтапную схему обработки изображений на основе анализа главных компонент [11, 12], и затем объединялись, используя алгоритм [7].

Подход 3. Исходные изображения сцен зашумлялись АБГШ с фиксированным σ , объединялись, используя алгоритм [7], и затем итоговое объединенное изображение подвергалось фильтрации, используя двухэтапную схему обработки изображений на основе анализа главных компонент [11, 12]. В данном случае предполагалось, что модель шума и его параметры на объединенном изображении являются идентичными тем, что были у исходных изображений до слияния.

В табл. 1 представлено численное сравнение различных подходов к объединению нескольких изображений фиксированной сцены для различных степеней зашумления σ . Наилучшие результаты выделены жирным шрифтом. Дополнительно на рис. 5 представлены визуальные результаты построения полностью сфокусированных изображений. Для проведения исследования использовались стандартные метрики эталонной оценки качества – ПОСШ и КСП. Данные метрики применялись к двум изображениям. Первое изображение, полученное с использованием алгоритма [7] в отсутствие шумов, бралось за эталонное. Второе – получено с использованием одного из трех вышеуказанных подходов.

Анализ результатов показывает, что качественное объединение зашумленных изображений фикси-

рованной сцены с использованием алгоритма [7] при наличии шумов возможно (рис. 5в), если исходные зашумленные данные будут предварительно обработаны с использованием определенной схемы фильтрации, например [11, 12]. Объединение изображений при отсутствии этой схемы не позволяет качественно идентифицировать пиксели, которые должны внести существенный вклад в построение полностью сфокусированного изображения. Как следствие, объединенное изображение в данном случае будет содержать шум, а также значительное размытие объектов интереса, присутствующих в сцене (рис. 5 б). Поэтому фильтрация после объединения не дает качественного результата обработки (рис. 5 г). Однако необходимо отметить, что фильтрация до объединения изображений приводит к многократному применению схемы шумоподавления. В этом случае каждое зашумленное изображение фиксированной сцены подвергается фильтрации, что значительно увеличивает вычислительную стоимость подхода 2. В случае подхода 3 фильтрация выполняется однократно применительно к объединенному изображению, что вычислительно менее затратно.

Таблица 1. ПОСШ, дБ / КСП полностью сфокусированных изображений при наличии АБГШ

σ	Подход 1	Подход 2	Подход 3
Сцена «Числовые карточки»			
5	31,28 / 0,749	32,51 / 0,876	34,83 / 0,879
15	24,21 / 0,416	31,99 / 0,862	31,83 / 0,845
20	22,03 / 0,308	31,82 / 0,857	31,25 / 0,838
25	20,26 / 0,230	31,80 / 0,852	30,84 / 0,833
35	17,52 / 0,140	31,46 / 0,843	30,17 / 0,825
Сцена «Робот»			
5	32,21 / 0,881	35,60 / 0,938	34,63 / 0,933
15	24,28 / 0,510	31,04 / 0,847	29,78 / 0,805
20	22,23 / 0,415	29,84 / 0,816	28,73 / 0,776
25	20,55 / 0,340	28,94 / 0,791	27,86 / 0,753
35	17,99 / 0,243	27,39 / 0,748	26,49 / 0,717
Сцена «Солдаты»			
5	31,87 / 0,907	33,85 / 0,954	33,63 / 0,953
15	23,32 / 0,584	28,39 / 0,869	26,42 / 0,809
20	21,34 / 0,488	27,06 / 0,834	24,94 / 0,752
25	19,79 / 0,412	26,04 / 0,801	23,92 / 0,709
35	17,40 / 0,298	24,53 / 0,742	22,55 / 0,646
Сцена «Игрушки»			
5	32,67 / 0,887	33,92 / 0,942	35,96 / 0,951
15	22,66 / 0,450	29,80 / 0,885	26,19 / 0,835
20	20,72 / 0,342	29,06 / 0,870	24,95 / 0,809
25	19,24 / 0,266	28,49 / 0,858	24,32 / 0,792
35	16,84 / 0,167	27,58 / 0,835	23,38 / 0,766

Заключение

Рассмотрены три варианта объединения зашумленных изображений: без предобработки и постобработки, с использованием предфильтрации объединяемых изображений, с использованием постфильтрации объединенного изображения. Анализ результатов показывает, что объединение зашумленных изображений с предфильтрацией позволяет более качественно идентифицировать пиксели, которые должны внести существенный вклад в построение полностью сфокусированного изображения. Однако необходимо отметить, что филь-

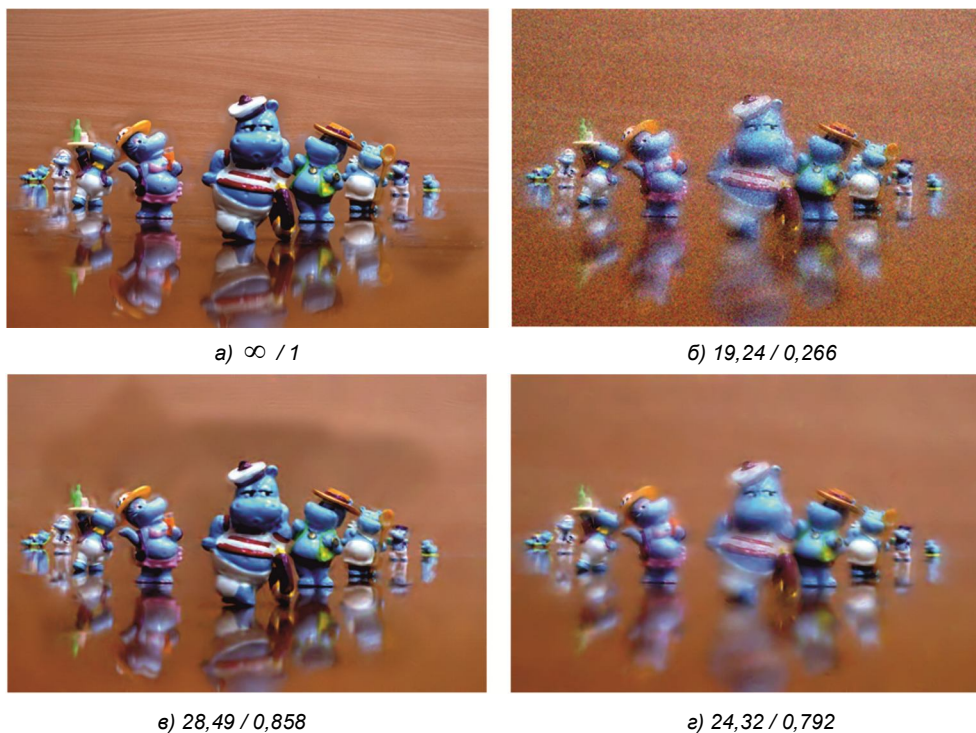


Рис. 5. Примеры построения полностью сфокусированных изображений для тестовой сцены «Игрушки» с указанием ПОСШ, дБ/КСП:

а) слияние без добавления АБГШ в объединяемые изображения;

б) слияние с добавлением АБГШ ($\sigma = 25$) в объединяемые изображения;

в) слияние с добавлением АБГШ ($\sigma = 25$) в объединяемые изображения и их фильтрацией до объединения;

г) слияние с добавлением АБГШ ($\sigma = 25$) в объединяемые изображения и фильтрацией объединенного изображения

рация до объединения изображений приводит к многократному применению схемы шумоподавления, что значительно увеличивает вычислительные затраты на обработку.

Работа выполнена при поддержке РФФИ (проект № 16-37-00301).

Литература

1. Гонсалес Р., Вудс Р. Цифровая обработка изображений. – М.: Техносфера, 2005.
2. Szeliski R. Computer vision: algorithms and applications. Springer, 2010.
3. Subbarao M., Choi T., Nikzad A. Focusing techniques // Optical Engineering J. 1993. V. 32, № 11. P. 2824–2836.
4. Shen C.-H., Chen H.H. Robust focus measure for low-contrast images // Proc. IEEE Int. Conf. Consumer Electronics. 2006. P. 69–70.
5. Li H., Chai Y., Yin H., Liu G. Multifocus image fusion and denoising scheme based on homogeneity similarity // Optics Communications J. 2012. V. 285, № 2. P. 91–100.
6. Naidu V.P.S., Raol J.R. Pixel-level image fusion using wavelets and principal component analysis // Defense Science J. 2008. V. 58, № 3. P. 338–352.
7. Носков А.А., Аминова Е.А., Приоров А.Л. Формирование полностью сфокусированных изображений на основе клеточных автоматов и пирамид изображений // Успехи современной радиоэлектроники. 2016. № 12. С. 39–46.
8. Katkovnik V., Foi A., Egiazarian K., Astola J. From

local kernel to nonlocal multiple-model image denoising // Int. J. Computer Vision. 2010. V. 86, № 8. P. 1–32.

9. Marsland S. Machine learning: an algorithmic perspective. Chapman and Hall, 2009.

10. Muresan D.D., Parks T.W. Adaptive principal components and image denoising // Proc. IEEE Int. Conf. Image Processing. 2003. V. 1. P. 101–104.

11. Волохов В.А., Сергеев Е.В., Приоров А.Л. Двух-этапная процедура фильтрации изображений на основе анализа главных компонент // Проектирование и технология электронных средств. 2011. № 2. P. 40–44.

12. Priorov A., Tumanov K., Volokhov V. Efficient denoising algorithms for intelligent recognition systems // In: Favorskaya M., Jain L.C. (eds.) Computer Vision in Control Systems-2, Intelligent Systems Reference Library, Springer International Publishing. 2015. V. 75. P. 251–276.

13. Сэлмон Д. Сжатие данных, изображений и звука. – М.: Техносфера, 2004.

14. Wang Z., Bovik A.C., Sheikh H.R., Simoncelli E.P. Image quality assessment: from error visibility to structural similarity // IEEE Trans. Image Processing. 2004. V. 13, № 4. P. 600–612.

15. Тоффолли Т., Марголус Н. Машины клеточных автоматов. – М.: Мир, 1991.

16. Hoekstra A.G., Kroc J., Sloot P.M.A. Simulating complex systems by cellular automata. Springer, 2010.

17. Burt P.J., Adelson E.H. The Laplacian pyramid as a compact image code // IEEE Trans. Communications. 1983. V. COM-31, № 4. P. 532–540.

К ВОПРОСУ О ВЛИЯНИИ СВЕТОЧУВСТВИТЕЛЬНОЙ МАТРИЦЫ НА КАЧЕСТВО ИЗОБРАЖЕНИЯ

Сорока Е.З., к.т.н., гл. научный сотрудник ЗАО «МНИТИ» (Московский научно-исследовательский телевизионный институт), e-mail: soroکا@mniti.ru.

ABOUT IMPACT OF PHOTO SENSOR MATRIX ON IMAGE QUALITY

Soroka E.Z.

Image distortions due to photo sensor matrix of TV camera have been analyzed by means of consecutive Fourier spectrum approach. Evaluations of arising linear and nonlinear distortion components have been defined. Method of sampling distortion reduction by means of postfiltering modification has been considered.

Key words: image distortion, photo sensor, sampling, Fourier spectrum, prefiltering, postfiltering.

Ключевые слова: искажения изображения, светочувствительная матрица, дискретизация, предфильтрация, постфильтрация.

Введение

Все датчики изображения, использующие для съемки светочувствительные матрицы, вносят в изображение определенные искажения, обуславливаемые двумерной пространственной фильтрацией и двумерной дискретизацией изображения.

Тракт передачи изображения в телевизионной или фотокамере содержит три последовательных звена: объектив, собственно матричный датчик и процессор сигналов (рис. 1).

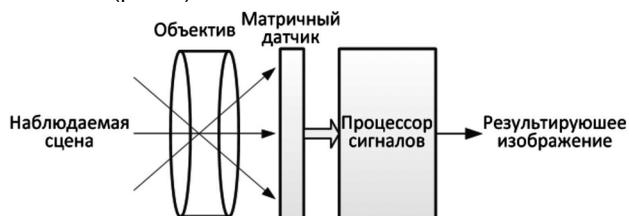


Рис. 1. Тракт передачи изображения

Действие объектива в первом приближении можно рассматривать как двумерную фильтрацию нижних частот, которая приводит к определенному ослаблению верхних частот в пространственном спектре формируемого изображения.

Действие матричного датчика описывается последовательно включенными фильтром низких частот и дискретизатором, которые более детально будут рассмотрены ниже.

Анализ дискретизации изображения светочувствительной матрицей

В целом, преобразования изображения в тракте передачи можно исследовать, используя классическую

Исследованы искажения изображения, вызываемые светочувствительной матрицей ТВ-камеры, с использованием последовательного спектрального метода. Получены оценки возникающих линейных и нелинейных составляющих искажений. Рассмотрен метод уменьшения искажений дискретизации путем модификации постфильтрации.

кую схему дискретизации, показанную на рис. 2 [1, 2].

Здесь использованы следующие обозначения:

$F(x, y)$, $F_1(x, y)$, $F_1^*(x, y)$, $F_2(x, y)$ – двумерные функции (изображения) соответственно на входе, после фильтра Φ_1 , после дискретизатора D и после фильтра Φ_2 ; $\mathcal{F}(f_x, f_y)$, $\mathcal{F}_1(f_x, f_y)$, $\mathcal{F}_1^*(f_x, f_y)$, $\mathcal{F}_2(f_x, f_y)$ – двумерные спектры соответствующих изображений; f_x и f_y – горизонтальная и вертикальная пространственные частоты; $G_1(x, y)$, $G_2(x, y)$ – импульсные характеристики фильтров Φ_1 и Φ_2 ; $\mathcal{K}_1(f_x, f_y)$, $\mathcal{K}_2(f_x, f_y)$ – частотные характеристики фильтров Φ_1 и Φ_2 ; $D(x, y)$ – двумерная дискретизирующая функция; $\mathcal{D}(f_x, f_y)$ – двумерная спектральная дискретизирующая функция.

Достоинство этой схемы состоит в том, что она позволяет четко разделить функции предфильтрации, постфильтрации и собственно дискретизации, которые в реальных технических системах практически всегда оказываются объединенными в конкретной аппаратуре.

Как известно, функция предфильтра Φ_1 реализуется за счет интегрирования входного излучения в пределах пиксельной площадки светочувствительной матрицы. В случае, если соседние площадки непосредственно граничат и имеют квадратную форму, двумерная импульсная характеристика предфильтра выражается произведением П-образных импульсных характеристик по обеим координатам в виде

$$G_1(x, y) = \Pi(nx) \Pi(ny), \quad (1)$$

где n – количество пикселей по вертикали (число строк), а $\Pi(z)$ определяется как

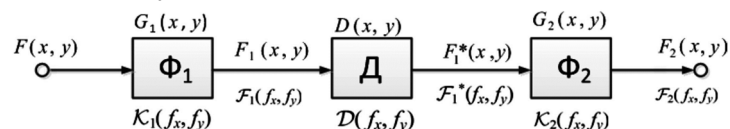


Рис. 2. Эквивалентная схема дискретизации изображения

$$\Pi(z) = \begin{cases} 1 & \text{при } |z| < \frac{1}{2}, \\ \frac{1}{2} & \text{при } |z| = \frac{1}{2}, \\ 0 & \text{при } |z| > \frac{1}{2}. \end{cases} \quad (2)$$

Здесь для унификации дальнейших оценок введена нормированная безразмерная высота изображения $h = 1$, что позволяет использовать единые шкалы частот в спектрах изображения независимо от его физического размера (в том числе и от размера матрицы). При этом унифицируется и размер пиксела, который можно записать как $1/n \times 1/n$. Кроме того, частота отсчетов (по обеим осям) составляет n , а максимальная частота, которая может быть передана отсчетами, составляет соответственно $n/2$.

Частотная характеристика фильтра Φ_1 выражается как преобразование Фурье от (1):

$$\mathcal{K}_1(f_x, f_y) = \frac{n}{2\pi f_x} \sin \frac{2\pi f_x}{n} \cdot \frac{n}{2\pi f_y} \sin \frac{2\pi f_y}{n}. \quad (3)$$

Эту частотную характеристику иллюстрируют линии уровня на рис. 3.

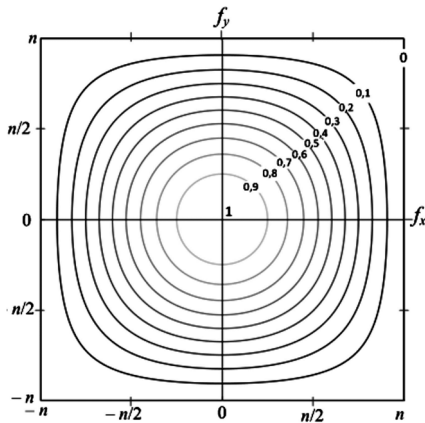


Рис. 3. Линии уровня частотной характеристики фильтра Φ_1

Светочувствительная матричная структура выполняет функции двумерного дискретизатора D , действие которого выражается матрицей отсчетов

$$F_1^*(x, y) = a^2 \sum_{k_1} \sum_{k_2} F_1(k_1 a, k_2 a) \delta(x - k_1 a) \delta(y - k_2 a), \quad (4)$$

где k_1 и k_2 – номера отсчетов по обеим осям, выражаемые целыми числами, δ – дельта-функция.

В результате дискретизации происходит периодизация спектра изображения, так что его спектр выражается в виде

$$F_1^*(f_x, f_y) = \sum_{l_1} \sum_{l_2} F_1(f_x - l_1 n, f_y - l_2 n), \quad (5)$$

где l_1 и l_2 – целые числа.

Это означает, что при дискретизации исходный спектр изображения периодизируется, т.е. повторяется со смещением по обеим частотным координатам с шагом, кратным n , т.е. числу строк изображения. При этом максимальная частота изображения, которая может быть воспроизведена без искажений в

соответствии с теоремой отсчетов Котельникова составляет $n/2$.

На рис. 4 представлен пример спектра изображения в виде двумерной гауссовой функции:

$$\mathcal{F}(f_x, f_y) = \exp[-a(f_x^2 + f_y^2)/4n^2], \quad (6)$$

где f_x, f_y – пространственные частоты в периодах на высоту изображения, n – число строк в изображении, a – постоянная, определяющая ширину спектра (т.е. скорость убывания спектра с частотой).

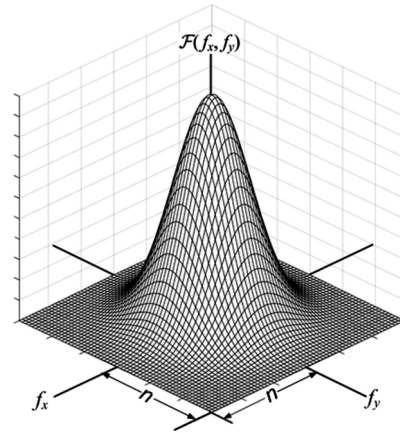


Рис. 4. Пример двумерного спектра изображения

На рис. 5 показан фрагмент спектра дискретизированного изображения $\mathcal{F}_1^*(f_x, f_y)$. Каждая составляющая спектра представлена двумя линиями уровня: зеленая линия соответствует границе спектра, при которой обеспечивается передача изображения без «ложных частот» – в соответствии с теоремой Котельникова, а черная линия уровня иллюстрирует пересечение составляющих при невыполнении теоремы.

Постоянная $a = 24$ подобрана таким образом, чтобы соседние по горизонтали и вертикали спектральные составляющие частично пересекались.

Таблица характерных значений спектра изображения

f	0	$n/2$	N
$\mathcal{F}_1(f)$	1	0,22	0,0025

Восстановление изображения из отсчетов выполняет фильтр Φ_2 , который должен подавлять побочные составляющие спектра. Как известно, функции этого фильтра выполняет устройство воспроизведения и зрение наблюдателя. Для упрощения двумерную пространственную частотную характеристику устройства воспроизведения можно считать совпадающей с частотной характеристикой предфильтра (3), т.е. принять

$$\mathcal{K}_2(f_x, f_y) = \frac{n}{2\pi f_x} \sin \frac{2\pi f_x}{n} \cdot \frac{n}{2\pi f_y} \sin \frac{2\pi f_y}{n}. \quad (7)$$

В результате спектр воспроизводимого изображения выражается как

$$\begin{aligned} \mathcal{F}_2(f_x, f_y) &= \mathcal{K}_2(f_x, f_y) \mathcal{F}_1^*(f_x, f_y) = \\ &= \mathcal{K}_2(f_x, f_y) \sum_{l_1} \sum_{l_2} F_1(f_x - l_1 n, f_y - l_2 n). \end{aligned} \quad (8)$$

Оценка искажений, вызываемых дискретизацией изображения

Если учитывать только четыре ближайших мешающие спектральные составляющие по f_x и f_y , показанные

на рис. 4, (8) сводится к следующему выражению:

$$\mathcal{F}_2(f_x, f_y) = \mathcal{K}_2(f_x, f_y) \cdot [\mathcal{F}_1(f_x, f_y) + \mathcal{F}_1(f_x - n, f_y) + \mathcal{F}_1(f_x + n, f_y) + \mathcal{F}_1(f_x, f_y - n) + \mathcal{F}_1(f_x, f_y + n)]. \quad (9)$$

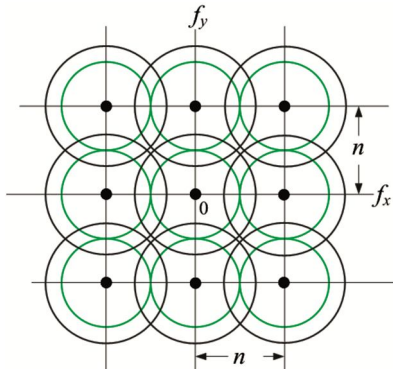
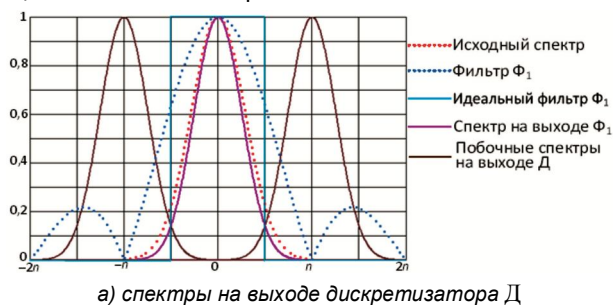


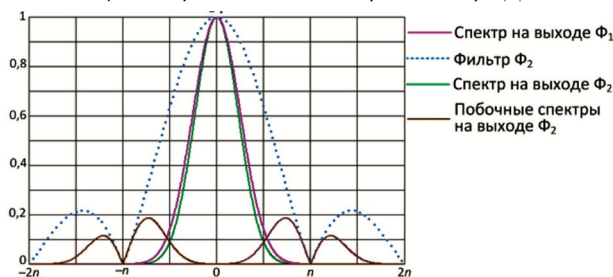
Рис. 5. Фрагмент спектра дискретизированного изображения

Для оценки возникающих искажений на рис. 6 приводятся графики сечений спектров по осям координат f_x и f_y .

Для оценки искажений, возникающих из-за мешающего действия четырех побочных составляющих спектра, расположенных по диагональным направлениям в плоскости частот (f_x, f_y) , на рис. 7 показаны графики сечений основного и мешающих спектров в этих направлениях в разных точках эквивалентной схемы дискретизации. Показан также вид частотной характеристики идеального фильтра, обеспечивающего выполнение теоремы отсчетов.



а) спектры на выходе дискретизатора Д



б) спектры на выхода фильтра Φ_2

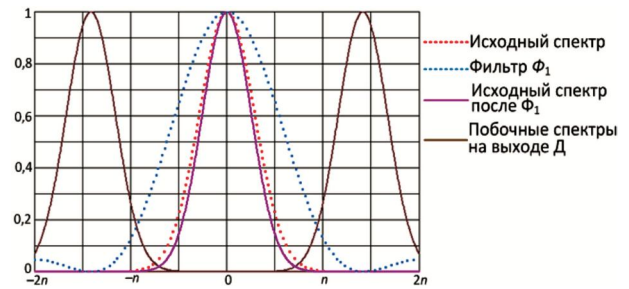
Рис. 6. Преобразования спектра изображения по частотам f_x и f_y ,

(частота в периодах на высоту изображения)

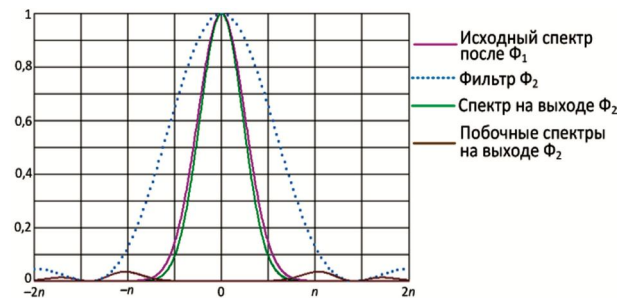
Как видно из сравнения мешающих спектров на выходе фильтра Φ_2 (коричневые графики) на рис. 6 и рис. 7, их проявление по «диагональным» частотам существенно ниже, чем по осям f_x и f_y .

Очевидными способами уменьшения искажений из-за дискретизации изображения светочувствительной матрицей являются улучшение характеристик

предфильтра Φ_1 и/или повышение частоты дискретизации путем перехода к матрице с кратно увеличенным числом пикселей по горизонтали и вертикали [3]. Однако как фильтр Φ_1 , так и дискретизатор Д в ТВ-камере имеют конструктивно фиксированные характеристики.



а) спектры на выходе дискретизатора Д



б) спектры на выходе фильтра Φ_2

Рис. 7. Преобразования спектра изображения по «диагональным» частотам (частота в периодах на высоту изображения)

Имеется простая возможность уменьшения искажений изображения путем модификации характеристик постфильтра Φ_2 . Эта модификация состоит в удвоении числа отсчетов по каждой координате, причем новые отсчеты получаются из имеющихся с помощью линейной интерполяции по каждой координате с помощью дискретного фильтра ([1], гл. 6) с импульсной характеристикой по оси x

$$g_d^*(x) = 0,5 \cdot \left(x + \frac{1}{2n} + \delta(x) + 0,5 \delta\left(x - \frac{1}{2n}\right)\right). \quad (10)$$

Здесь размер пиксела (шаг дискретизации в матрице) по-прежнему составляет $1/n$.

Благодаря такой фильтрации число отсчетов удваивается, и новые отсчеты находятся как среднее двух соседних исходных. Частотная характеристика этого фильтра имеет вид

$$k_d^*(f) = \cos^2(\pi f / 2n). \quad (11)$$

Следует заметить, что переход от линейной интерполяции к более сложным алгоритмам дает лишь незначительное улучшение [5].

Завершающим элементом модифицированного постфильтра является аналоговый фильтр с частотной характеристикой (определяемой с учетом уменьшения вдвое размера пиксела)

$$k_a(f) = \text{sinc}(\pi f / 2n), \quad (12)$$

который соответствует структуре матрицы средства отображения.

Применение подобного фильтра для преобразования по оси y обеспечивает получение новых пикселей из исходных путем билинейной интерполяции [4].

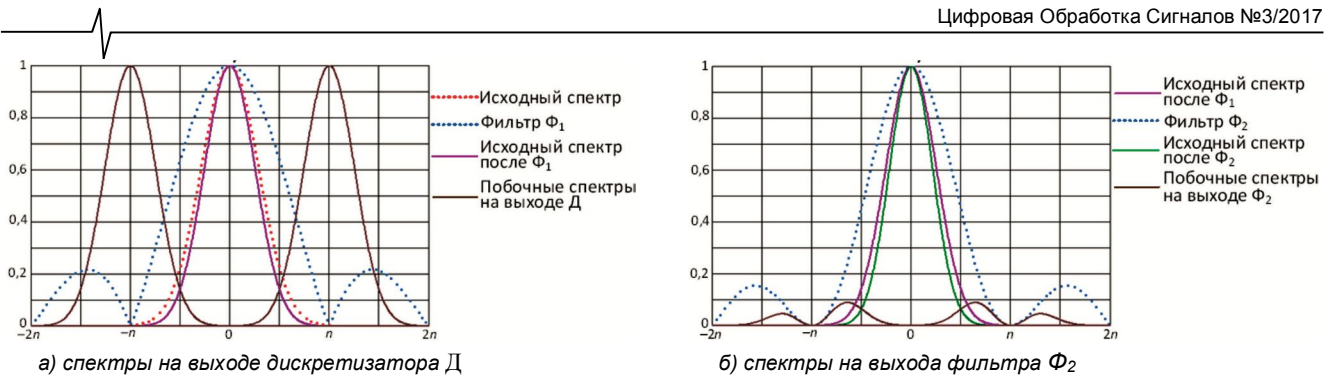


Рис. 8. Преобразования спектра изображения

по частотам f_x и f_y при использовании интерполяции пикселей (частота в периодах на высоту изображения)



а)



б)

в)

Рис. 9. Иллюстрация обработки изображения с удвоением числа пикселей по каждой координате и интерполяцией

На рис. 8 показаны графики сечений основного и побочных спектров в разных точках эквивалентной схемы дискретизации с использованием постфильтра, состоящего из дискретного фильтра (11), обеспечивающего интерполяцию отсчетов, и выходного аналогового фильтра с частотной характеристикой (12).

Как видно из сравнения спектральных характеристик на рис. 6(б) и 8(б), применение интерполяции пикселей дает существенное уменьшение амплитуды побочных составляющих спектра (не менее 6 дБ). Тем самым обеспечивается заметное повышение качества результирующего изображения. Это иллюстрирует рис. 9, где показан женский портрет (а), увеличенный фрагмент портрета (б) и этот же фрагмент с удвоенным количеством пикселей по каждой оси и интерполяцией. В последнем фрагменте практически полностью устранены артефакты дискретизации – блочность (мозаичность) изображения и ступенчатость наклонных линий.

Заключение

Выполнен детальный анализ искажений изображения, вызываемых светочувствительной матрицей

ТВ-камеры, с использованием спектрального подхода. Получены выражения для оценки искажений, вызываемых двумерной предфильтрацией в матрице, собственно дискретизацией за счет матрицы и двумерной постфильтрацией при восстановлении изображения.

Исследован вариант постфильтрации с использованием удвоения числа пикселей по каждой координате и получением дополнительных пикселей путем интерполяции исходных, что обеспечивает существенное ослабление искажений, вызываемых дискретизацией – мозаичности изображения и ступенчатости наклонных линий.

Литература

1. Игнатьев Н.К. К дискретизации сигналов с неограниченным спектром. – Электросвязь, 1960, № 2, с. 71-72.
2. Игнатьев Н.К. Дискретизация и ее приложения. – М.: Связь, 1980. – 264 с.
3. Красильников Н.Н. Цифровая обработка 2D- и 3D-изображений: учеб. пособие. – СПб.: БХВ-Петербург, 2011. – 608 с. – гл. 4.
4. Прэйтт У. Цифровая обработка изображений. Кн. 1. – М.: Мир. – 1982 – 312 с. – Гл. 4.
5. Ваганов С.Е., Хашин С.И. Сравнение алгоритмов удвоения размеров изображения. // Моделирование и анализ информационных систем – 2016. – Т. 23. – № 4. – С. 389–400.

КРАТКИЕ СООБЩЕНИЯ

УДК681.324.687

РАЗРАБОТКА АРХИТЕКТУРЫ КМОП-АЦП С ШИРОКИМ ДИАПАЗОНОМ ЧАСТОТ ПРЕОБРАЗОВАНИЯ ОТ НЕСКОЛЬКИХ КИЛОГЕРЦ ДО НЕСКОЛЬКИХ ГИГАГЕРЦ

Кононов В.С., к.т.н., АО «Специализированное конструкторско-технологическое бюро электронных систем», e-mail: casandra1983@mail.ru.

DESIGN OF ARCHITECTURE OF CMOS-ADC WITH WIDE BANDWIDTH FROM FEW KILOHERTZ TO FEW GIGAHERTZ

Kononov V.S.

Key words: architecture, design, CMOS-ADC, capacitor, bandwidth, range.

Ключевые слова: АЦП, конвейер, конденсатор, частота, преобразование.

Введение

В настоящее время основную долю рынка высокоскоростных аналого-цифровых преобразователей (АЦП) занимают многоразрядные преобразователи с конвейерной архитектурой [1] (рис. 1).

Такая архитектура отличается способностью одной секции обрабатывать данные другой секции в пределах одного такта. В результате выходной сигнал предыдущей секции передается в следующую секцию через устройство выборки-хранения (УВХ), а секция получает новые данные для дальнейшего преобразования. В этом состоит принцип конвейерного преобразования. Следует отметить, что благодаря УВХ конвейерные АЦП имеют широкий динамический диапазон, так как УВХ нивелирует неидентичность задержек компараторов внутри каждой секции.



Рис. 1. Блок-схема типичной

К-разрядной секции конвейерного АЦП: $AC_{вх}$,

$AC_{вых}$ – аналоговые сигналы на входе и выходе секции;

ЦС – цифровой сигнал на выходе секции

Роль УВХ, как известно, заключается в способности сохранять информацию о сигнале на входе секции до момента завершения, так называемого, «грубого» преобразования, что позволяет вычислить сигнал остатка на выходе секции и передать его в следующую секцию для дальнейшего преобразования.

Однако эта способность не гарантирует сохранение необходимой точности в течение продолжительного времени. Дело в том, что конденсатор, обычно используемый в качестве элемента хранения (рис. 2), теряет заряд под воздействием различного рода утечек, основными из которых являются подпороговый ток и утечки сток/источниковых $p\text{-}$ переходов ключевого МОП-транзистора [2].

Рассмотрены особенности классических КМОП-АЦП с конденсаторной архитектурой с точки зрения зависимости полосы частот преобразования от емкости конденсаторов хранения промежуточной информации. Показано, что для расширения полосы частот в сторону низких значений можно использовать встроенный генератор тактовых сигналов с высокой фиксированной частотой, обеспечивающей максимально возможную частоту преобразования и накопления цифровых данных в выходном регистре АЦП, вывод которых становится возможным практически на любой частоте внешнего источника тактовых сигналов.

В результате стремление к повышению частоты преобразования, которое обычно сопровождается уменьшением емкости конденсатора хранения, неминуемо приводит к повышению минимальной частоты преобразования, что ограничивает использование высокоскоростных КМОП-АЦП на низких частотах.

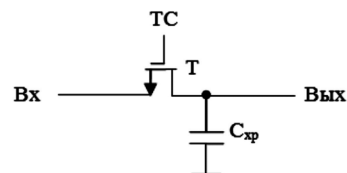


Рис. 2. Устройство выборки-хранения:

T – ключевого транзистор; $C_{хр}$ – конденсатор хранения;

ТС – тактовый сигнал

Целью данной статьи является разработка архитектуры КМОП-АЦП на основе переключаемых конденсаторов, позволяющей осуществлять преобразование сигналов одним АЦП в широком диапазоне частот от нескольких килогерц до нескольких гигагерц.

Направивается техническое решение, основанное на использовании встроенного генератора тактовых сигналов с высокой фиксированной частотой, при которой обеспечивается максимально-возможная скорость преобразования (рис. 3).

В результате проблема изменения зарядового состояния конденсатора хранения устраняется, а вывод цифровой информации из выходного регистра становится возможным практически на любой частоте, не превышающей частоту встроенного генератора.

Такое решение фактически превращает обычный

КМОП-АЦП с классической конденсаторной архитектурой в АЦП со свойствами, присущими бесконденсаторной архитектуре [3, 4]. В дальнейшем для удобства анализа такой КМОП-АЦП будем называть псевдо бесконденсаторным АЦП.

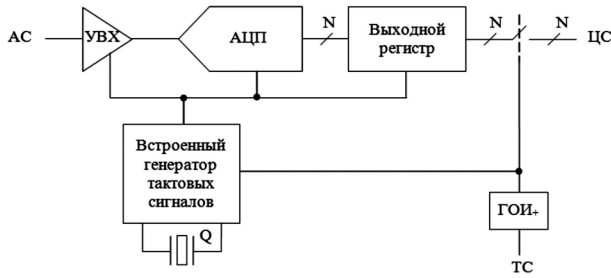


Рис. 3. Блок-схема АЦП со встроенным генератором тактовых сигналов с фиксированной частотой: ГОИ+ – генератор одиночных импульсов, возникающих по переднему фронту ТС; Q – кварц

Практическое применение псевдо бесконденсаторных АЦП имеет свои особенности. Дело в том, что любой генератор тактовых сигналов, как внешний, так и внутренний в АЦП, должен быть «кварцованным», чтобы удовлетворять определенным требованиям по «дрожанию» тактовых сигналов [1]. Из-за существующих ограничений по доступной номенклатуре кварцев в АЦП с частотой преобразования до нескольких гигагерц не удастся воспользоваться простой схемой непосредственной генерации тактовых сигналов с требуемой частотой (рис. 4).

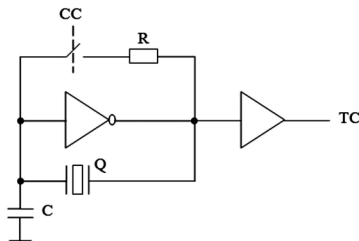


Рис. 4. Схема простого кварцевого генератора тактовых сигналов: R, C – элементы калибровки; СС – сигнал синхронизации

Поэтому на кристалле КМОП-АЦП необходимо размещать встроенный генератор с петлей фазовой автоподстройки основной частоты (ФАПЧ), в которой схема (рис. 4) будет играть роль источника опорной частоты. Так как опорная частота является фиксированной и более низкой по сравнению с основной ча-

стотой, то обеспечить устойчивую работу встроенного генератора и в целом петли ФАПЧ не представляет большой сложности.

Выводы

Проведенный анализ позволяет сделать следующие выводы:

1. Классические конденсаторные КМОП-АЦП с конвейерной архитектурой широко представлены на рынке, но имеют ограничения по полосе тактовых частот не только сверху, но и снизу. Ограничения сверху определяются в основном инерционностью транзисторов и паразитными емкостями соединительных шин на кристалле. Ограничения снизу связаны с изменением зарядового состояния переключаемых конденсаторов из-за подпороговых токов ключевых МОП-транзисторов и утечек стоков/истоковых $p\bar{n}$ -переходов этих транзисторов.

2. Использование встроенных генераторов тактовых сигналов высокой фиксированной частотой, обеспечивающей максимально возможную скорость преобразования, позволяет снять проблему изменения зарядового состояния конденсаторов хранения и, одновременно, создать возможность вывода цифровой информации из выходного регистра АЦП практически на любой частоте, не превышающей частоту встроенного генератора.

3. Для выполнения требований по «дрожанию» тактовых сигналов встроенный генератор должен быть «кварцованным» и охваченным петлей ФАПЧ, необходимость которой обусловлена не схемотехнической целесообразностью, а ограничениями по номенклатуре кварцев, доступных для применения в России.

Литература

1. Кестер У. Аналого-цифровое преобразование [Текст]: монография / У. Кестер; пер. с англ. Е.Б. Володина. – М.: Техносфера, 2007. – 1016 с.
2. Rudy van de Plassche. CMOS integrated analog-to-digital and digital-to-analog converters [Text] / Rudy van de Plassche. – Kluwer academic publishers, 2003. – 588 p.
3. Кононов В.С. Архитектуры бесконденсаторных КМОП-КНИ-АЦП для космического применения [Текст] / В.С. Кононов // Вестник Воронежского государственного технического университета. – 2016. – Т. 12. – №1, С. 70-73.
4. Кононов В.С. Особенности вычисления и обработки сигналов остатка в бесконденсаторных АЦП с секционированной архитектурой [Текст] / В.С. Кононов // Цифровая обработка сигналов. – 2017.–№1, С. 70-72.

Уважаемые коллеги!

Для тех, кто не успел оформить подписку на второе полугодие 2017 года через АО «Роспечать», сохраняется возможность приобретения журналов непосредственно в редакции по адресу: г. Москва, ул. Авиамоторная, дом 8, Научный Центр МТУСИ, ком. 612. Российское научно-техническое общество радиотехники, электроники и связи им. А.С. Попова, метро «Авиамоторная», или оформить Заказ в соответствии с требованиями, выставленными на сайте журнала: www.dsps.ru.

Справки по телефону: (+7 903)201-53-33 (Самсонов Геннадий Андреевич).

E-mail: rntores@mail.ru



AUTEX Ltd.



IEEE



20-я Международная Конференция

ЦИФРОВАЯ ОБРАБОТКА СИГНАЛОВ И ЕЁ ПРИМЕНЕНИЕ

Москва, ИПУ РАН, март 2018 года

*The 19th International Conference DIGITAL SIGNAL
PROCESSING AND ITS APPLICATIONS*

Уважаемые коллеги!

ПРИГЛАШАЕМ ВАС ПРИНЯТЬ УЧАСТИЕ В РАБОТЕ КОНФЕРЕНЦИИ

ОРГАНИЗАТОРЫ:

- Российское научно-техническое общество радиотехники, электроники и связи им. А.С. Попова
- Институт радиотехники и электроники им. В.А. Котельникова РАН
- Компания AUTEX Ltd.
- Российская секция IEEE
- IEEE Signal Processing Society
- Институт проблем управления РАН
- Институт проблем передачи информации РАН
- Московский научно-исследовательский телевизионный институт (ЗАО «МНИТИ»)

ПРИ УЧАСТИИ:

- Федеральное агентство по промышленности РФ
- Министерство образования и науки РФ
- Институт проблем информатики РАН
- ОАО «РТИ»
- ОАО «ГСКБ «АЛМАЗ-АНТЕЙ»
- ЗАО «Инструментальные системы»
- НТЦ «МОДУЛЬ»
- ЗАО «СКАН Инжиниринг Телеком»
- ОАО «Концерн «СОЗВЕЗДИЕ»
- ГУП НПЦ «ЭЛВИС»
- Владимирский государственный университет
- Московский физико-технический институт (университет)
- Московский авиационный институт (университет)
- Московский государственный технический университет им. Н.Э. Баумана
- Московский институт радиотехники, электроники и автоматизации (университет)
- Московский институт электронной техники (университет)
- Московский технический университет связи и информатики
- Московский энергетический институт (университет)
- Рязанский государственный радиотехнический университет
- Санкт-Петербургский государственный университет телекоммуникаций им. проф. А.М. Бонч-Бруевича
- Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина)
- Ульяновский государственный технический университет
- Ярославский государственный университет

МЕЖДУНАРОДНЫЙ ОРГАНИЗАЦИОННЫЙ КОМИТЕТ КОНФЕРЕНЦИИ:

Председатель Международного комитета: академик РАН Гуляев Юрий Васильевич

Сопредседатели: академик РАН Васильев С.Н., академик РАН Каляев И.А., Кирпичников А.П., академик РАН Сигов А.С., академик РАН Федоров И.Б.

Программная группа Международного оргкомитета:

Председатель Программной группы: член-корр. РАН Зубарев Юрий Борисович

Члены Программной группы: проф. Аджемов А.С. (Россия), член-корр. РАН Борисов В.И. (Россия), д.т.н. Борисов Ю.И. (Россия), академик РАН Бугаев А.С. (Россия), проф. Витязев В.В., проф. Лукин В.В. (Украина), проф. Митра М. (США), проф. Петровский А.А. (Белоруссия), проф. Сарамаки Т. (Финляндия), член-корр. РАН Черепенин В.А. (Россия).

Административная группа Международного оргкомитета: проф. Бартенев В.Г., проф. Брюханов Ю.А., проф. Васильев К.К., проф. Дворкович В.П., проф. Дворкович А.В., проф. Джиган В.И., Зубарев Р.В., проф. Пожидаев В.Н., д.т.н. Приоров А.Л., проф. Прохоров Ю.Н., доц. Самсонов Г.А., проф. Санников В.Г., доц. Сергиенко А.Б., проф. Скачко Ю.В., проф. Солонина А.И., проф. Соколов И.А., проф. Сперанский В.С., доц. Штенченко В.Б., проф. Фаворская М.Н., проф. Чобану М.К., проф. Шинаков Ю.С., проф. Ямпурин Н.П.

РАБОТА КОНФЕРЕНЦИИ ПЛАНИРУЕТСЯ ПО СЛЕДУЮЩИМ СЕКЦИЯМ:

1. Теория сигналов и систем
2. Теория и методы ЦОС
3. Обработка сигналов в системах телекоммуникаций
4. Обработка сигналов в радиотехнических системах
5. Обработка и передача изображений
6. Обработка и передача измерительной информации
7. Проектирование и техническая реализация систем ЦОС
8. Системы и устройства «технического зрения»
9. Цифровая обработка многомерных сигналов
10. Нейрокомпьютерная обработка сигналов и изображений