# ЦИФРОВАЯ ОБРАБОТКА СИГНАЛОВ

Научно-технический журнал № 3/2020

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., член-корр. РАН А.В. 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.И. Джиган, д.т.н., проф. В.В. Еремеев, д.т.н. Г.В. Зайцев, Р.В. Зубарев, А.П. Кирпичников, д.т.н., акад. РАН Н.А. Кузнецов, В.Г. Мистюков, д.т.н., проф. С.Л. Мишенков, д.т.н., проф. Ю.Н. Прохоров, д.т.н., проф. Ю.Н. Прохоров, д.т.н., проф. Чиров Д.С., к.т.н., проф. В.С. Сперанский.

#### 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 30.09.20 г. Формат 60х90/8.

Гарнитура «Arial». Печать офсетная. Бумага офсетная. Печ.л. 6,5. Тираж 200 экз.

Заказ № 6115. 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.

#### УЧРЕДИТЕЛИ:

# АО «Инструментальные системы» ФГУП «НИИ радио»

ООО «Российское научно-техническое общество радиотехники, электроники и связи им. А.С. Попова»

#### B HOMEPE:

Кузнецов А.Е., Кочергин А.М. Современные подходы по организации обработки и обеспечения потребителей данными ДЗЗ высокого разрешения	3
Мятов Г.Н., Юдаков А.А., Кузнецов П.К., Мартемьянов Б.В., Еремеев В.В., Кузнецов А.Е.	
Мониторинг состояния бортовых систем космических аппаратов оптического наблюдения Земли в обеспечение автоматической наземной обработки получаемых изображений	9
Ларионов С.М., Макаренков А.А.	
Кластеризация гиперспектральных спутниковых снимков земной поверхности на основе метода плотности ближайших соседей	15
Кузнецов А.Е., Рыжиков А.С.	
Программное управление многосерверной обработкой изображений большого объема от систем космического наблюдения Земли	21
Лаврушкин С.В., Ватолин Д.С.	
Разработка нейросетевых методов распространения цвета в видео	25
Стефаниди А.Ф., Приоров А.Л., Топников А.И., Хрящев В.В.	
Модификация VGG-архитектуры в задачах унимодальной и мультимодальной биометрии	35
Трофимов Д.В.	
Повышение качества сопровождения объектов на основе анализа минимумов критериальной функции	41
Никишкин П.Б., Витязев В.В.	
Методы широкополосной передачи данных на основе сигналов с частотным разделением каналов	45
Пономарев А.В., Пономарева О.В.	
Роль и место операции дополнения нулями в теории двумерной Фурье-обработки сигналов	50

Журнал «Цифровая обработка сигналов» включен в перечень ведущих рецензируемых научных изданий, в которых по рекомендации минобрнауки РФ, должны быть опубликованы основные научные результаты диссертаций на соискание ученой степени доктора и кандидата наук

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

# Digital Signal Processing

Science & Technical Magazine Issue 3, 2020 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 Corresponding Member of RAS A.V. 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
Djigan V.I., Dr. Sci (Tech), Professor
Eremeyev V.V., Dr.Sci.(Tech.), Professor
Zaitsev G.V., Dr.Sci.(Tech.)
Zubarev R.V.,
Kirpichnikov A.P.,

Kuznetsov N.A., Dr.Sci.(Tech.), Academician Mistyukov V.G., Mishenkov S.L., Dr.Sci.(Tech.), Professor Priorov A.L., Dr. Sci (Tech)

Prokhorov Y.N., Dr.Sci.(Tech.), Professor Sannikov V.G., Dr.Sci.(Tech.), Professor Chirov D.S., Dr.Sci.(Tech.), Professor Speranskii V.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.dspa.ru

### Publisher:

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

## Computer makeup:

I.A. Blagodarova

### FOUNDERS: InSys

# Radio Research and Development Institute

#### **CONTENTS:**

Kuznetcov A.E., Kochergin A.M., Laryukov S.A.  Modern approaches to organization of processing and providingconsumers with high-resolution remote sensing data3
Myatov G.N., Udakov A.A., Kuznechov P.K., Martemiyanov B.V., Eremeev V.V., Kuznechov A.E.
Monitoring of technical conditions of on-board optical  Earth observation satellite systems in support of automatical ground-based image processing9
Larionov S.M., Makarenkov A.A.
Clustering of hyperspectral satillite images of the earth's surface based on the nearest neighbors density method15
Kuznechov A.E., Rijikov A.S.
Multiserver processing software management
for large-scale remote sensing imagery21
Lavrushkin S.V., Vatolin D.S.
Development of neural networks for color propagation in video25
Stefanidi A. F., Priorov A. L.,
Topnikov A. I., Khryashchev V. V.
Modification of VGG-architecture for unimodal
and multimodal biometrics35
Trofimov D.V.
Improving the quality of object tracking based on the analysis
of minima of the criterion function41
Nikishkin P.B., Vityazev V.V.
Broadband data transmission methods based on signals
with frequency-division multiplexing45
Ponomarev A.V., Ponomareva O.V.
The role of zero padding in the theory of two-dimensional
Fourier signal processing50



Subscription index: 82185 ("Rospechat") ISSN: 1684-2634 9 771684 263005

# СОВРЕМЕННЫЕ ПОДХОДЫ ПО ОРГАНИЗАЦИИ ОБРАБОТКИ И ОБЕСПЕЧЕНИЯ ПОТРЕБИТЕЛЕЙ ДАННЫМИ ДЗЗ ВЫСОКОГО РАЗРЕШЕНИЯ

Кузнецов А.Е., д.т.н., первый зам. директор, НИИ «Фотон» РГРТУ, e-mail: foton@rsreu.ru; Кочергин А.М., к.т.н., с.н.с. НИИ «Фотон» РГРТУ, e-mail: foton@rsreu.ru; Ларюков С.А., инженер НИИ «Фотон» РГРТУ, e-mail: foton@rsreu.ru.

## MODERN APPROACHES TO ORGANIZATION OF PROCESSING AND PROVIDING CONSUMERS WITH HIGH-RESOLUTION REMOTE SENSING DATA

#### Kuznetcov A.E., Kochergin A.M., Laryukov S.A.

The paper describes the technical approaches that allow to perform a monitoring and control of processing both operational and archive remote sensing information, to organize remote interaction of consumers with data processing center using the cloud technologies of data processing and storage.

Key words: Earth remote sensing, geoportal, browser, map server, WebSocket, cloud service, space monitoring.

**Ключевые слова:** дистанционное зондирование Земли, геопортал, браузер, картографический сервер, веб-сокет, облачный сервис, космический мониторинг.

Предлагаются технические решения, позволяющие контролировать и управлять процессом обработки как оперативной, так и ретроспективной информации ДЗЗ, организовать удаленное взаимодействие потребителей с ЦОД на основе облачных технологий обработки и хранения данных.

#### Введение

В отличие от глобальных и региональных космических систем наблюдения системы дистанционного зондирования Земли (ДЗЗ) высокого пространственного разрешения предоставляет потребителям изображения небольших по площади территорий. С учетом этой особенности различают два режима получения спутниковой информации. Первый — это оперативный режим, в соответствии с которым по заявкам потребителей планируется съемка требуемых территорий, затем осуществляется прием и обработка данных по одному из стандартных уровней, а затем передача результатов заказчику. Ключевым элементом этого режима является оперативное получение потребителем интересующих материалов.

Второй режим – ретроспективный, он связан с анализом и отбором архивных изображений, хранящихся на центре обработки данных (ЦОД), формировании заявки на их обработку и последующее получение выходных информационных продуктов. Для этого режима важным является наличие многофункционального сервиса по анализу архивных данных, а также возможности получения продуктов с минимальными временными задержками.

Для обеспечения потребителей данными Д33 созданы: комплекс программно-технических средств взаимодействия с потребителями, каталог материалов съемки и WEB-средства для их просмотра, средства автоматизированного формирования выходных информационных продуктов. При создании этих средств использовались технические решения 10-15-летней давности, что приводит к существенным временным задержкам при обеспечении потребителей данными Д33. Например, сервис по просмотру и отбору материалов съемки не предоставляет возможность анализировать детальные изображения земной поверхности, распространение инфор-

мации осуществляется через FTP-сервер или с помощью оптических дисков, а не в режиме Online [1, 2, 3]. Для визуального анализа выходных информационных продуктов потребителю необходимо иметь специализированное программное обеспечение, как правило зарубежного производства. Отсутствие возможности со стороны заказчика контролировать процесс исполнения заявок делает крайне неудобным взаимодействие поставщика и потребителя спутниковой информации.

Большинство из имеющихся проблем могут быть успешно решены, если автоматизировать процесс управления обработкой информации и обслуживания потребителей с возможностью современных серверных WEB-технологий. В настоящей работе предлагаются технические решения, позволяющие контролировать и управлять процессом обработки как оперативной, так и ретроспективной информации, организовать удаленное взаимодействие потребителей с ЦОД на основе облачных технологий обработки и хранения данных, что в конечном итоге повысит доступность информации ДЗЗ для массовых пользователей.

#### Функциональная схема ЦОД

На рис. 1 представлена функциональная схема составных частей ЦОД, отвечающих за обеспечение потребителей спутниковой информацией.

Источником данных ЦОД являются маршруты съемки наблюдаемых территорий, поступающих от антенного комплекса приема информации (оперативный режим) или выбираемых из архива центра. Получение выходных информационных продуктов осуществляется в комплексе обработки по заданиям, формируемым комплексом контроля и управления. Для оперативного обслуживания потребителей через WEB-портал, растровые изображе-

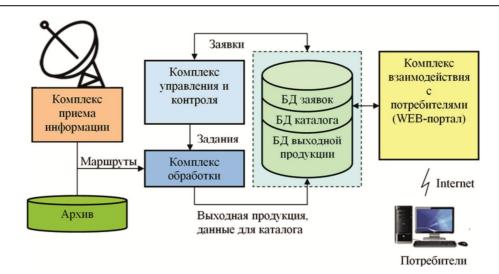


Рис. 1. Функциональная схема взаимодействия составных частей ЦОД

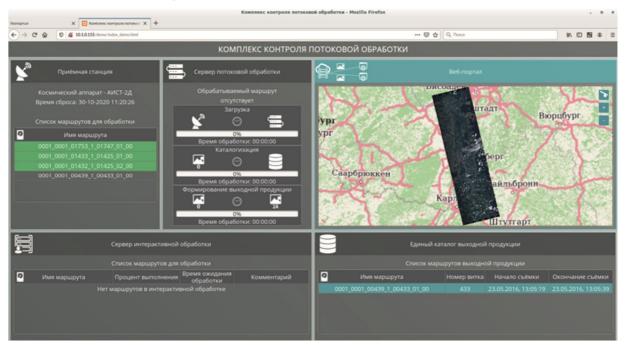


Рис. 2. Главное окно комплекса управления и контроля

ния тейлируются и сохраняются в файловой базе данных (БД). Для маршрутов, полученных от комплекса приема, рассчитываются метаданные для каталога, содержащие квиклуки и характеристики видеоинформации. По данным из каталога с помощью WEB-портала осуществляется поиск интересующей информации в архиве центра.

Задания на обработку видеоданных формируются на основе заявок потребителей через интерфейс WEB-портала.

На комплекс контроля и управления, помимо формирования заданий на обработку, отводится важная роль по контролю реализации всех технологических операций, начиная от приема данных до размещения результатов обработки в БД геопортала (БД выходной продукции). Этот комплекс предоставляет информацию о работоспособности всех подсистем ЦОД, функционирующих в автоматическом режиме, и принятия своевременных мер в случае возникновения нештатных ситуаций. На рис. 2 показаны инструментальные панели

комплекса при его штатной работе, а именно:

- панель контроля работы приемной станции;
- панель маршрутов, поступающих на сервер потоковой обработки;
- панель с информацией о размещении обрабатываемых маршрутов в БД геопортала и каталоге.

Далее рассмотрим принципы построения и функционирования WEB-портала комплекса взаимодействия с потребителем.

#### Структура геопортала

Геопортал проектируется по клиент-серверной архитектуре [4], которая включает в себя серверное ПО, установленное на Интернет-сервере, и клиентское ПО, выполняемое в браузере пользователя.

Серверное ПО выполняет функции регистрации и аутентификации пользователей, реализации поиска спутниковых изображений в архиве по заданным параметрам, формирования географической подложки на основе векторных картографических данных и предо-

ставления интерфейса создания заявок на новую съёмку и обработку архивных снимков.

Клиентское ПО предназначено для организации интерфейса по взаимодействию пользователя с серверным ПО. Его основными функциями являются: организация интерфейса регистрации и аутентификации пользователей, предоставление интерфейса ввода параметров для поиска спутниковых изображений в электронном каталоге, формирования запросов на новую съёмку и обработку архивных снимков, вывод на странице пользователя результатов поиска в табличном и графическом видах, уведомление клиента об изменении статуса созданных им заявок, обеспечение их визуального просмотра и скачивания результатов, отображение полученной с сервера многослойной картографической основы в графическом виде в браузере клиента, визуализация спутниковых данных на фоне картографической основы, организация мониторинга территорий, предоставление инструментов обработки растровых изображений, нанесение векторных объектов на картографическую основу.

Разрабатываемый геопортал базируется на свободно-распространяемых программах и библиотеках. Часть функций серверного ПО выполняют картографический сервер GeoServer [5] и сервер WebSocket Swoole [6]. Функции клиентского ПО по работе с картографической основой выполняет картографическая библиотека OpenLayers [7].

Исходя из представленных функций и решения об использовании специального картографического ПО, разработана структурная схема геопортала, общий вид которого представлен на рис. 3.

Для повышения производительности и обеспечения безопасности серверное ПО разделено на четыре блока, каждый из которых функционирует на отдельной аппаратной базе.

- 1. Сервер оперативного архива.
- 2. Сервер системы управления базами данных (СУБД).
  - 3. Веб-сервер.
  - 4. Сервер комплекса обработки.

На сервере оперативного архива хранятся спутни-

ковые снимки высокого пространственного разрешения и БД архива спутниковых снимков. Спутниковые снимки занимают большой объём дискового пространства (как правило, более 10 ГБ) и предоставляются пользователям после предварительной обработки. БД архива спутниковых снимков наполняется в результате каталогизации данных ДЗЗ и содержит в себе атрибутивную информацию по условным кадрам маршрута съёмки и обзорные снимки (квиклуки) кадров.

Сервер СУБД – это компьютер с развёрнутой на нём системой PostgreSQL [8], содержащей в своём составе следующие БД: копии БД архивов спутниковых снимков, БД пользователей геопортала, БД векторной картографической информации.

Копии БД архива спутниковых снимков представляют собой доступные только для чтения частичнореплицированные копии БД архива спутниковых снимков. Использование копий исключает прямой доступ к «рабочим БД» со стороны серверного ПО геопортала. Сервер СУБД геопортала позволяет подключать несколько БД архива спутниковых снимков, содержащих в себе разнотипную информацию, полученную из различных источников.

БД пользователей геопортала содержит в себе общие сведения о клиентах, созданных ими регионах интереса, информацию по сформированным заявкам на новую съёмку, заданиям на обработку архивных данных, а также данные по группам пользователей.

БД векторной картографической информации является основой для формирования географической подложки. Источником данных для нее стали подробные и бесплатно распространяемые карты мира Open-StreetMap [9].

Веб-сервер – Интернет-сервер, на котором реализована логическая часть серверного ПО. В состав веб-сервера входит программа геопортала, сервер Web-Socket, тайловый сервер и GeoServer.

Программа геопортала реализует логику серверной части системы и отвечает за формирование вебинтерфейса в браузере пользователя, авторизацию, регистрацию, разграничение прав пользователей и взаимодействие с БД.

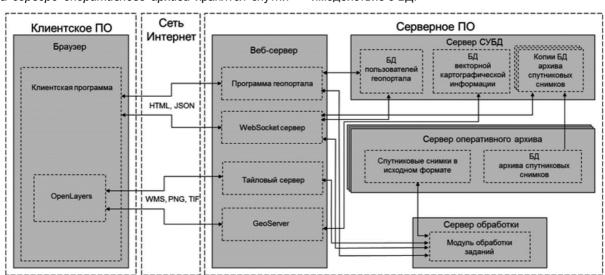


Рис. 3. Структурная схема геопортала

GeoServer выполняет следующие функции: приём запросов и передача ответной информации клиентам по протоколу WMS [10], растеризация векторной картографической информации, реализация методов по разграничению доступа пользователей к геоданным, обеспечение ввода, отображение и редактирование картографических данных пользователей.

Тайловый сервер осуществляет обработку запросов с клиентской части на выдачу растровых данных детальных снимков.

Сервер WebSocket Swoole предоставляет поддержку технологии WebSocket, необходимую для оперативного обмена данными между серверной и клиентской частью геопортала.

#### Функциональные возможности

Одной из основных функций геопортала является поиск архивных данных в электронном каталоге. Для предоставления пользователю единого интерфейса поиска спутниковых изображений, полученных от сканеров различного типа и хранящихся в различных БД в геопортале, реализуется динамическая система поиска. Ее принцип заключается в следующем: для каждой БД архивных данных разрабатывается специальный модульпровайдер, обеспечивающий унифицированный доступ к БД с учетом ее структуры. После регистрации провайдера в системе геопортала в пользовательском интерфейсе появляется возможность осуществлять поиск архивных спутниковых снимков, информация о которых хранится в соответствующей БД.

При осуществлении поиска в каталоге геопортал собирает информацию о найденных спутниковых снимках посредством вызова функций поиска каждого из под-

ключенных провайдеров, составляет сводную таблицу атрибутивной информации и отправляет результат в браузер пользователя. По мере готовности результаты поиска в виде растровых изображений последовательно отправляются в клиентскую программу геопортала при помощи технологии WebSocket, где наносятся на картографическую основу.

На рис. 4 показан результат поиска видеоданных в трех различных БД, содержащих информацию о восьми космических аппаратах и соответствующих им съёмочных устройствах, имеющих различный принцип построения изображений.

Другой важной функцией геопортала является инициализация удаленной обработки архивных данных ДЗЗ. Для получения обработанных спутниковых снимков пользователю необходимо составить задание на обработку. После составления задание отправляется на вебсервер и сохраняется в БД пользователей геопортала. Пользователь может наблюдать за ходом выполнения задания на обработку в личном кабинете. Модуль обработки заданий сервера обработки отслеживает появление новых записей в БД пользователей геопортала. При появлении такой записи сервер обработки обращается к серверу оперативного архива для получения спутниковых снимков. После этого запускается автоматическая обработка согласно параметрам задания. Полученный результат сервис обработки данных отправляет в геопортал, где результирующее изображение подготавливается для онлайн просмотра, а также сжимается в ZIPархив для предоставления возможности его скачивания пользователем. Одновременно в личном кабинете пользователя появляется уведомление о выполнении задания

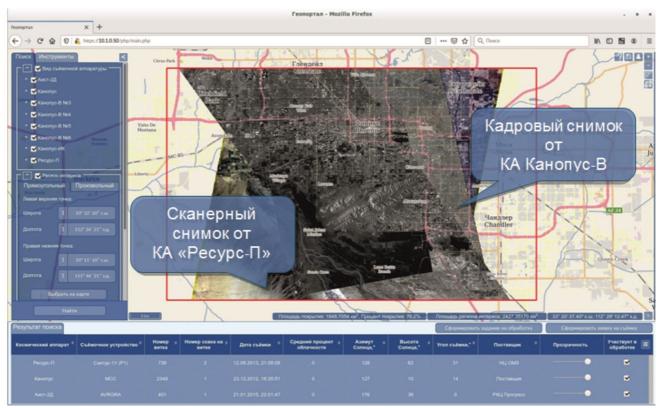


Рис. 4. Отображение видеоданных, полученных с различных космических аппаратов и съемочных устройств

Отображение картосновы и обработанных высокодетальных космических снимков в браузере осуществляется в картографической проекции Web Mercator и построено на базе тайловой архитектуры. Поверхность Земли условно разбивается на тайловую сетку при различных уровнях масштаба. Клиентское ПО запрашивает у серверной части тайлы в формате PNG, которые согласно тайловой сетке и масштабу должны быть отображены и выводит их на экран. Соответственно, для отображения высокодетальных изображений необходимо организовать процесс, который будет по запросу с клиентской части разбивать исходный спутниковый снимок на тайлы в заданных проекции и формате для дальнейшего отображения в браузере пользователя.

Используемый картографический сервер GeoServer позволяет организовать такой процесс, однако он не может эффективно работать с изображениями формата GeoTIFF размером более двух гигабайт [14]. Для работы со спутниковыми снимками, размер которых превышает данное значение, необходимо построить пирамиду [15], представляющую из себя несколько копий мозаик изображения с разным уровнем масштабирования. Каждый элемент мозаики называется тайлом и хранится как отдельное изображение. Для того чтобы GeoServer смог получить доступ к пирамиде, тайлы необходимо записать в БД в формате PostGIS Raster [16].

При получении запроса от клиентского ПО GeoServer преобразует параметры тайловой сетки в картографические координаты и осуществляет поиск в БД тайлов пирамиды. Результатами поиска являются тайлы пирамиды в формате GeoTIFF. Далее GeoServer выполняет трансформирование тайлов пирамиды из проекции детального снимка в проекцию Web Mercator, изменяет формат изображения с GeoTIFF на PNG, нарезает его в соответствии с параметрами тайловой сетки и отправляет в клиентскую часть.

Такой подход обладает универсальностью, однако имеет недостатки, заключающиеся в многократном преобразовании формата изображения и трансформировании системы координат, что приводит к серьезному ухудшению производительности системы. Для того, чтобы избежать этот негативный эффект было решено не использовать в процессе отображения высокодетальных снимков GeoServer, а реализовать свой тайловый сервер. Исходя из того, что картографическая проекция и соответствующая ей тайловая сетка фиксированы, в нем был реализован следующий техпроцесс визуализации данных.

Исходный снимок трансформируется в проекцию Web Mercator с дальнейшим построением пирамиды. Каждый слой пирамиды разделяется на тайлы формата TIFF размером 256 на 256 пикселей и заносится в БД согласно тайловой сетке. Использование формата TIFF позволяет сохранить в тайле исходную глубину цвета, а также все каналы спутникового изображения.

При поступлении в тайловый сервер запроса тайла с клиентской части он выполняет поиск нужного изображения в БД с использованием не картографических координат, а номеров строк, столбцов и уровней приближения тайловой сетки, что значительно сокращает вре-

мя поиска и нагрузку на сервер БД. Одновременно происходит выборка и загрузка атрибутивных параметров спутникового снимка. В клиентской части осуществляется декодирование полученного с сервера тайла из формата ТІFF, выбор комбинации каналов изображения и преобразование тайла в поддерживаемый браузером формат PNG.

Применение собственного тайлового сервера для отображения высокодетальных спутниковых снимков позволило существенно увеличить производительность системы, а также предоставило возможность выполнять обработку нанесённых данных в клиентской программе геопортала без дополнительного обращения к серверу, поскольку в браузер пользователя загружаются атрибутивные данные снимка и информация обо всех каналах изображения с исходной глубиной цвета.

Предложенный механизм позволил реализовать в геопортале динамическую систему обработки изображений, основанную на модульном принципе. Для реализации новой функции обработки требуется создать модуль (фильтр) обработки и зарегистрировать его в системе. После этого его интерфейс автоматически станет доступен пользователю. Перед отрисовкой тайла на экране менеджер фильтров передаёт указатель на данный тайл каждому из зарегистрированных модулей, которые выполняют соответствующую обработку растра с использованием имеющейся атрибутивной информации. В результате на экран выводится тайл с наложенными на него фильтрами.

Предложенный подход сокращает нагрузку на сервер за счёт перемещения части обработки видеоданных на сторону клиента. Это позволяет избежать повторной передачи подвергшихся обработке тайлов с сервера на клиент, тем самым уменьшив потребляемый системой Интернет-трафик и существенно повысив скорость визуализации. Кроме того появляется возможность развивать функциональность геопортала без внесения изменений в основную программу.

В современных условиях динамично изменяющегося мира все более важной становится задача организации мониторинга участков земной поверхности. Анализ нескольких существующих геоинформационных сервисов, в том числе «Банк Базовых продуктов» и «LandViewer» [11] показал, что в них не реализован автоматический мониторинг территорий, а отслеживание региона интереса возможно только в ручном режиме.

В разрабатываемом геопортале мониторинг территорий реализуется в автоматическом режиме, то есть пользователь указывает регион поиска и критерии интересующих изображений, а система в автоматическом режиме осуществляет поиск спутниковых снимков, удовлетворяющих заданным критериям и отображает их в браузере. Таким образом, пользователю доступна актуальная информация без дополнительных действий с его стороны.

При реализации данного механизма возникла проблема в организации максимально оперативной передачи данных с сервера на рабочее место клиента по мере появления новой информации без увеличения нагрузки на сервер и объема передаваемых данных.

Традиционным способом обмена данными между браузером и серверным ПО является подход, называемый AJAX («асинхронный JavaScript и XML»).

АЈАХ – подход к построению интерактивных пользовательских интерфейсов веб-приложений, заключающийся в «фоновом» обмене данными браузера с вебсервером [12].

При применении технологии АЈАХ для оперативного обмена данными между клиентом и сервером используются методы частого опроса сервера (Polling) и очереди ожидающих запросов (Long Poll). Метод Polling отличается простотой реализации, однако имеет недостатки, заключающиеся в периодичном формировании запросов с клиента и постоянной нагрузке на сервер. Метод Long Poll лишён недостатков метода Polling, однако оба метода построены по принципу, когда инициатором обмена является клиентская сторона. Однако в режиме автоматического мониторинга инициатором передачи данных является сервер, что не позволяет использовать технологию AJAX без увеличения нагрузки на сервер и канал передачи данных. Вместо нее решено использовать технологию WebSocket [13].

Протокол WebSocket способствует распространению интерактивного содержимого и созданию приложений реального времени, а также реализует возможность обмена данными между браузером и сервером через постоянное соединение по инициативе серверной части. Данные передаются по нему в обоих направлениях в виде «пакетов», без разрыва соединения и дополнительных HTTP-запросов. Технология WebSocket обеспечивает достаточную безопасность за счёт использования шифрования и существенное сокращение сетевого трафика.

Использование веб-сервера с поддержкой протокола WebSocket позволит в разрабатываемом геопортале предоставить максимально оперативный доступ к спутниковым снимкам, принятым на приёмной станции и обработанным на сервере обработки в режиме реального времени.

#### Заключение

По результатам проведённых исследований были созданы действующие макеты комплекса контроля и комплекса взаимодействия с потребителями (Webпортал). Отработка комплекса управления происходила на данных от КА «Ресурс-П». Для взаимодействия с комплексом обработки был разработан протокол информационного взаимодействия, позволяющий задавать требования по получению выходной продукции уровней обработки 1 и 2, а также диагностировать сам процесс обработки.

Основной задачей при испытаниях Web-портала была оценка его производительности. Результаты тестирования Web-сервера на базе 4-ядерного процессора с часто-

той 3,4 ГГц и 16 ГБ ОЗУ показали, что для 100 одновременно обращающихся пользователей среднее время на получение ответа составило 450 мс при 60 % загрузке процессора. На основе полученных оценок были сформулированы требования к серверу и накопителю БД выходной продукции, использование которых в Web-портале позволяет хранить ≈ 1,2 ГБ видеоданных и обслуживать запросы пользователей с задержкой не более 1 с.

Практическое использование созданного программного обеспечения планируется в рамках центров обработки данных от КА «Аист-2Т» и от 4-х малых космических аппаратов сверхвысокого разрешения, создаваемых АО «Корпорация «ВНИИЭМ».

#### Литература

- 1. Онлайн-портал Европейского Космического Агенства ESA [Электронный ресурс]. URL: https://scihub.co-pernicus.eu/dhus/#/home/.
- 2. ArcGIS [Электронный ресурс]. URL: https://www.es-ri.com/ru-ru/arcgis/products/arcgis-online/overview
- 3. Банк базовых продуктов. Руководство пользователя [Электронный ресурс]. URL: https://bbp.ntsomz.ru/assets-landing/BBP Handbook 20201008.pdf.
- 4. D. Jacobson, G. Brail, D. Woods. APIs: A Strategy Guide. O'Reilly Media, Inc. 2001. 150 p.
- 5. Документация Geoserver [Электронный ресурс]. URL http://docs.geoserver.org/latest/en/docguide.
- 6. Документация Swoole [Электронный ресурс]. URL: https://www.swoole.co.uk/docs/.
- 7. OpenLayers [Электронный ресурс] URL: http://openlayers.org/.
- 8. A.Ю. Васильев Работа с PostgreSQL: настройка и масштабирование. Creative Commons Attribution-Noncommercial 4.0 International. 2017. 288 с.
- 9. Экспорт OpenStreetMap [Электронный ресурс] URL: https://wiki.openstreetmap.org/wiki/RU:Экспорт/.
- 10. Сервисы доставки данных OWS [Электронный ресурс] URL: http://gis-lab.info/qa/ows.html.
- 11. EOS LandViewer [Электронный ресурс]. URL: https://eos.com/landviewer/.
- 12. Б. Маклафлин. Изучаем Ajax = Head Rush Ajax. СПб.: Питер, 2007.
- 13. Спецификация RFC 6455 [Электронный ресурс]. URL: https://tools.ietf.org/html/rfc6455
- 14. Raster Data Optimization [Электронный ресурс]. URL: https://geoserver.geo-solutions.it/edu/en/enterprise/raster.html.
- 15. Building and using an image pyramid [Электронный pecypc]. URL: https://docs.geoserver.org/latest/en/user/tutorials/imagepyramid/ima gepyramid.html.

Storing a coverage in a JDBC database [Электронный ресурс]. – URL: https://docs.geoserver.org/stable/en/user/tutorials/imagemosaic-jdbc/imagemosaic-jdbc\_tutorial.html.

8

УДК 528.855

# МОНИТОРИНГ СОСТОЯНИЯ БОРТОВЫХ СИСТЕМ КОСМИЧЕСКИХ АППАРАТОВ ОПТИЧЕСКОГО НАБЛЮДЕНИЯ ЗЕМЛИ В ОБЕСПЕЧЕНИЕ АВТОМАТИЧЕСКОЙ НАЗЕМНОЙ ОБРАБОТКИ ПОЛУЧАЕМЫХ ИЗОБРАЖЕНИЙ

Мятов Г.Н., д.т.н., зам. главного конструктора АО «РКЦ «Прогресс», e-mail: miatov@mail.ru

Юдаков А.А., к.т.н., начальник отдела АО «РКЦ «Прогресс»

Кузнецов П.К., д.т.н., зав. отделом, Самарский федеральный исследовательский центр РАН, e-mail: kurnesov@mail.ru

Мартемьянов Б.В., к.т.н., доц., СамГТУ, e-mail: bvmart@rambler.ru

Еремеев В.В., д.т.н., директор НИИ «Фотон» РГРТУ, e-mail: foton@rsreu.ru

Кузнецов А.Е., д.т.н., зам. директора НИИ «Фотон» РГРТУ

# MONITORING OF TECHNICAL CONDITIONS OF ON-BOARD OPTICAL EARTH OBSERVATION SATELLITE SYSTEMS IN SUPPORT OF AUTOMATICAL GROUND-BASED IMAGE PROCESSING

Myatov G.N., Udakov A.A., Kuznechov P.K., Martemiyanov B.V., Eremeev V.V., Kuznechov A.E.

Within the framework of activities aimed at increasing the competitiveness of domestic species products of remote sensing of the Earth, an integrated approach to the diagnosis of on-board systems of optical-electronic satellite (OES) is proposed for consideration in order to determine the suitability of information obtained by the OES spacecraft for the automatic creation of high-quality geoformation products.

**Key words:** monitoring, remote sensing of the Earth, diagnosis of on-board systems, high-quality geoformation products.

**Ключевые слова:** дистанционное зондирование Земли, оптико-электронный телескопический комплекс, наземная обработка информации, линейное разрешение, точность геодезической привязки.

Введение

Сегодня трудно найти отрасль народного хозяйства, в которой не использовалась бы видеоинформация, получаемая спутниками ДЗЗ. Область применения этой информации расширяется по мере улучшения характеристик доставляемых на Землю снимков по разрешению, точности определения географических координат наблюдаемых объектов, а также с появлением бортовых мультиспектральных и гиперспектральных фотопреобразователей. Новейшие оптико-электронные КА получают видеоданные с разрешением на местности, достигающих долей метра, с точностью определения координат объектов порядка единиц метров. Поставщики видеоинформации, в первую очередь разработчик космической системы (КС) ДЗЗ, и ее потребители заинтересованы в получении исходной информации ДЗЗ с гарантированными показателями качества. Для этого в рамках наземной инфраструктуры КС создается система предварительной обработки видеоинформации, позволяющая проводить оценку ее качества, а также диагностировать состояние бортовых систем КА [1-3], входящих в контур получения и первичного преобразования видеоданных. Такой подход позволяет на ранней стадии процесса обработки и передачи видеоданных сохранять только ту часть, которая пригодна для создания в автоматическом режиме качественных геоинформационных

В рамках деятельности, направленной на повышение конкурентоспособности отечественной видовой продукции дистанционного зондирования Земли (ДЗЗ), предлагается к рассмотрению комплексный подход к диагностике бортовых систем оптико-электронных космических аппаратов (КА) в целях определения пригодности информации, поступающей с КА ДЗЗ, для создания в автоматическом режиме качественных геоинформационных продуктов.

продуктов, например, высокоточных геопространственных данных.

В ряде работ, посвященных исследованию высокодетальных оптоэлектронных автоматических КА ДЗЗ показана связь качества получаемой сигнальной и координатной информации об объектах наблюдаемой сцены с эффективностью работы бортовых систем КА, обеспечивающих функционирование целевой аппаратуры [4, 5].

Для контроля качества функционирования КА на орбите необходимо в составе наземного комплекса обработки информации реализовать комплексную диагностику состояния бортовых навигационных систем и оптико-электронного телескопического комплекса (ОЭТК). Такая диагностика должна быть ориентирована на использование комплексной информации, поступающей с бортовых приборов КА: сигнальной, координатной, телеметрической, а также видеоданных.

Методология диагностирования качества функционирования КА базируется на двух методах. Первый метод направлен на уточнение и диагностирование состояния элементов внутреннего и внешнего ориентирования КА. Метод позволяет системно увязать процесс геометрической полетной калибровки КА по выявлению пространственного и углового положения визирной оси съемочной

аппаратуры относительно приборной системы координат звездных датчиков по тестовым наземным полигонам и процесс наземной обработки бортовой измерительной информации (БИИ) и видеоинформации (ВИ).

Схема уточнения и диагностирования состояния элементов внутреннего и внешнего ориентирования КА приведена на рис. 1, на котором введены следующие дополнительные обозначения: НА-ВРЛ — наземная аппаратура высокоскоростной радиолинии; БССИ — бортовая система сбора измерительной информации; НССИ — наземная система сбора измерительной информации; РП КА — рабочая программа для космического аппарата; СА — съемочная аппаратура; РП ВА — рабочая программа для специальной аппаратуры; ПСП — план сеансов приема.

В обеспечение наземной обработки БИИ в целях достижения требуемого качества геопространственной продукции на борту КА с помощью БССИ проводится получение исходных измерений, характеризующих угловое и линейное положение КА, а также угловую скорость КА. Собранная БИИ передается в сеансе связи с использованием НА-ВРЛ, после чего она выделяется из основного потока информации и поступает в комплекс обработки БИИ, где проводится уточнение и синтез элементов внутреннего и внешнего ориентирования. Результаты обработки представляют собой файл с метаданными, использование которых обеспечивает требуемые фотограмметрические характеристики получаемых высоко детальных снимков.

Геометрическая полетная калибровка по одноименным точкам, проводимая для определения параметров внутреннего и внешнего ориентирования КА, реализуется на основе метода самокалибровки [6-8]. Периодическое проведение процедуры калибровки позволяет определять углы ориентации приборных систем координат звездных датчиков относительно визирной системы координат съёмочной аппаратуры с ошибкой менее одной угловой секунды, а погрешность определения положения ПЗС-матриц в фокальной плоскости, с учётом

дисторсии и фокусного расстояния, с ошибкой, не превышающей одного пикселя.

Калибровка каждой точки изображения включает измерения пиксельных координат изображения «эталонных» (калибруемых) объектов местности, геодезические координаты которых известны на двух и более изображениях [9-12]. Проводимые операции калибровки позволяют определить значения элементов внешнего и внутреннего ориентирования КА с точностями, которые обеспечивает требуемое качество геоинформационных продуктов и на которые следует ориентироваться в процессе диагностирования состояния бортовых систем КА. Файл метаданных, содержащий сопроводительную информацию к обрабатываемому снимку с указанной точностью, обеспечивает возможность автоматической обработки снимков.

Таким образом, использование «эталонных» значений параметров ориентации КА на момент съемки поверхности Земли позволяет реализовать фильтрацию входных данных, поступающих с борта КА в сеансе приема информации, и проводить высокоточную обработку изображений в автоматическом режиме. На основе изложенного подхода разработан и реализован в космической системе «Ресурс-П» программный продукт по автоматической обработке изображений в наземном комплексе приема, обработки и распространения информации (НКПОР-РП).

Опыт эксплуатации КА «Ресурс-П» № 1, 2, 3 показал, что разработанная компоновочная схема и элементы конструкции КА обладают высокой динамической стабильностью на момент съемки поверхности Земли. Динамическая стабильность конструкции оценивается на уровне «шумовой» работы звездных датчиков и составляет менее 3 угловых секунд. В связи с этим возникает возможность повысить точность геодезической привязки снимков путем установки более точных навигационных приборов и доработки наземного программного обеспечения. Такая доработка, помимо учета особенностей сбора и синтеза информации, поступающей с навига-

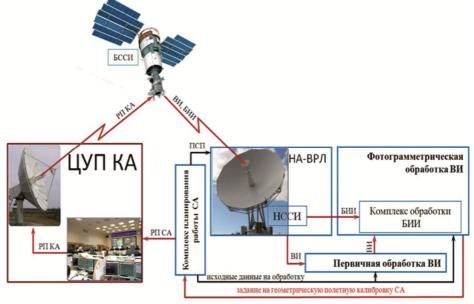


Рис. 1. Схема уточнения и диагностирования элементов внутреннего и внешнего ориентирования

ционных приборов, касается, в том числе, и совершенствования методики обработки калибровочных данных. В частности, геометрическая полетная калибровка целевой аппаратуры КА с датчиками сканового принципа действия, размещаемыми в фокальной плоскости в «шахматном» порядке, имеет ряд особенностей, которые заметно проявляются при возрастании требований к точности привязки. В процессе регистрации последовательности изображений с подвижного основания, которым является КА, возникают непрерывные изменение ракурса съемки, что приводит к появлению заметных (на субпиксельном уровне) проективных различий в обрабатываемых строках двух, полученных последовательно во времени кадров изображений одной и той же местности. Кроме того, в процессе формирования наборов видеоданных из-за дискретного передвижения зарядовых пакетов по столбцам ПЗС матриц, работающих в режиме временной задержки и накопления (ВЗН), а также вибраций бортовой аппаратуры вызванных дисбалансами вращающихся роторов силовых гироскопов либо маховиков, возникает «смаз» видеоданных. Проективные геометрические искажения и «смаз» видеоданных затрудняют автоматическое совмещение фрагментов изображений. Чем более высокие требования предъявляются к точности координатной привязки снимков, тем более существенным становится влияние рассогласования проективных параметров совмещаемых фрагментов снимка, а также эффекта смаза изображения.

Помимо указанных причин, на рассогласование координат сопряженных точек на наблюдаемых и эталонных изображениях влияют как случайные, так и систематические ошибки измерений углового и линейного положения КА в процессе съёмки, разделить которые можно только обрабатывая достаточно большую серию снимков.

В связи с этим, создание эталонов геопривязанных изображений осуществляется идентификацией одноименных точек не на единичных снимках, а на серии снимков наземных полигонов, объекты на которых имеют известные координаты геопривязки. Идентификация одноименных точек изображений, точнее говоря малых фрагментов, накрывающих эти точки, производится поиском и высокоточным совмещением их с имеющимися изображениями «опорных» точек, географические координаты которых известны. При обработке серии изображений опорных полигонов должно быть идентифицировано достаточно большое количество опорных точек местности с известными геодезическими координатами и высотой. Точность совмещения изображений при использовании взаимно корреляционной функции в значительной степени зависит от ее крутизны в окрестности экстремума. Поэтому при использовании корреляционного метода при нечетко выраженном экстремуме [9], например, для фрагментов изображений типа облаков, полей, вызывает определенные затруднения. В результате поле совмещенных фрагментов для подобных сюжетов получается сильно разреженным, что не всегда удовлетворяет требованиям практических задач. Ниже рассматривается более точный метод совмещения изображений, названный методом функционализации, который в определенной степени, свободен от этого недостатка [13-16].

# Высокоточное совмещение изображений – метод функционализации

Будем считать, что ракурсные искажения достаточно малы, чтобы можно было совмещать изображения простым плоским смещением одного изображения относительно другого. При таком условии рассмотрим методику решения практически важной задачи идентификации параметров вибраций фокальной плоскости ОЭТК и значений параметров вектора смаза получаемых изображений на примере данных от КА типа «Ресурс-П». Информация о значениях указанных параметров позволит точнее производить апостериорное вычисление кинематических параметров КА и, соответственно, улучшить точность геопривязки космических изображений.

Процедура совмещения изображений использует специфику конструкции оптико-электронного преобразователя (ОЭП) видеоаппаратуры КА. ОЭП составлен из размещенных на фокальной плоскости в шахматном порядке нескольких десятков линейных фоточувствительных матриц (ПЗС линеек), поля зрения которых на краях матриц пересекаются на 2-3 десятка пикселов. Изображения в области пересечения полей зрения пзсматриц создают последовательность двух кадров с периодом, равным времени перемещения («пролета») точек изображений между ПЗС-матрицами. Время пролета оценивается с использованием процедуры совмещения таких пар кадров. Причем, время пролета в процессе съемки маршрута может непрерывно и достаточно быстро меняться, что повышает требования к быстродействию и точности алгоритмов совмещения. Указанных двукадровых последовательностей изображений получается достаточно много, всего на единицу меньше количества пзс-матриц, составляющих ОЭП. Это дает возможность соответствующей обработкой повысить достоверность оценок параметров относительного сдвига изображений и вычислять для фиксированного момента времени эпюр распределения этих параметров вдоль ОЭП. По этой информации, с использованием экспертных оценок, возможно выявить источники возмущений и принять конструктивные или алгоритмические меры по их нейтрализации.

Предлагаемая методика совмещения изображений апробирована при решении задачи определения параметров «смаза» изображения [13, 14], причиной появления которого являлась несинхронность движения изображения по фокальной плоскости ОЭТК и движения зарядовых пакетов по столбцам ПЗС матриц, составляющих ОЭП. Методика базируются на так называемом методе функционализации параметров изображений [15, 16]. Метод функционализации, в отличие от широко используемых в аналогичных задачах экстремальнокорреляционных и градиентных методов, позволяет достаточно просто реализовать быстрые итерационные процедуры высокоточного совмещения изображений. В основе метода лежит соотношение, устанавливающее связь идентифицируемых параметров покадрового сдвига изображений с их измеримыми характеристиками. Под измеримой характеристикой фрагмента изображения понимается усредненное некоторым образом значение освещенности этого фрагмента.

Суть метода функционалиации следующая. Анализируются два полученных последовательно во времени изображения подвижной сцены. Обозначим их как  $E_0$ ,  $E_1$ :  $E_0=E(x,t_0),\ E_1=E(x,t_1),$ 

Необходимо найти преобразование:  $E_0 \to E_1$ . При этом наложим ограничение на допустимое преобразование: будем считать, что изображение  $E_1$  образуется плоско-параллельным синхронным смещением всех точек изображения  $E_0$  с параметрами вектора смещения, одинаковыми для всех его точек. В таком случае можно записать:

$$\begin{split} E_1(x,t) &= E_0(x-s(t)), \\ \frac{ds}{dt} &= v(t), \quad s(t_0) = 0, \, 0 = (0 \, 0), \end{split} \tag{1}$$

где v = (v, u) – вектор скорости движения изображения (так называемый оптический поток);  $t \in R^1$ ; v, u – соответственно проекции вектора скорости v на оси Ox и Oy;

 $s=(s_x,\ s_y)$  — соответственно проекции на оси Ox и Oy вектора сдвига s изображения, происшедшего на интервале времени [  $t_0,t_1$ ]. Соотношение (1) называется уравнением движения изображения [15].

В большинстве работ, например [9], посвященных анализу движения изображения, для вычисления оценок оптического потока используются операция дифференцирования. Но реальные изображения в подавляющем большинстве случаев представляются функциями распределения яркостей, имеющими разрывы первого рода. Такие разрывы имеют место, например, на переходах свет/тень, крыша здания/земля, автомобиль/дорожное покрытие и т.п. Поэтому перед дифференцированием изображение подвергается операции сглаживания, например, гаусс-фильтром. Этот прием делает дифференцирование корректным, но «уничтожает» мелкие, часто очень важные, детали в изображениях и, соответственно, приводит к локальным ошибкам в получаемых оценках оптического потока. С целью преодоления этого недостатка в [14, 15, 16] предложено совмещать не сами изображения, а индуцируемые ими поля значения функционала специального вида F(E(t)), обладающего свойством дифференцируемости. Функционал F(E(t)) строится как комбинация регулярных обобщенных функций, определенных на совмещаемых изображениях как

$$F(E(t)) = \int_{D} K(x) E(x,t) dxdy,$$
 (2)

где функция веса K(x) функционала стеснена следующими условиями: а)  $K(x)\colon R2 \Rightarrow R1;$  б)  $\forall\, x\in \Gamma(D)\, K(x)=0;$  D- область определения функционала (далее, окно анализа),  $\Gamma-$  граница без особенностей окна D; в) непрерывна, ограничена и дифференцируема почти всюду по всем своим аргументам;  $\Gamma$   $\int_D K(x,y)\, dxdy=0$ .

Свойство г) обеспечивает нулевое среднее поля значений функционала. В рассматриваемой задаче равенство нулю среднего значения отсчетов изображений значительно упрощает процедуру совмещения.

Функционал (2) называется основным функционалом излагаемого здесь метода функционализации. Основной

функционал корректно дифференцируем по всем своим аргументам, поскольку дифференцирование выражения (2) по пространственным координатам сводится [17] к дифференцированию функции веса K(x), которая по определению дифференцируема по всем своим аргументам.

Очевидно, что функционал (2) обладает фильтрующими свойствами и его применение сглаживает изображение, но, как известно [17], сжатием области определения D функционала всегда можно достичь желаемого компромисса между точностью дифференцирования и степенью сглаживания.

Рассмотрим сначала случай непрерывного движения изображения. Для получения соотношения, связывающего вектор скорости движения динамического изображения с его измеримыми характеристиками, вычислим полную производную функционала (2) по времени, учитывая, что его значения изменяются только в силу уравнения движения изображения (1) (условие постоянства яркости динамического изображения):

$$\dot{F}(E(t)) = F_x(E(t)) v + F_y(E(t)) u,$$
 (3)

где 
$$\dot{F} = \frac{dF}{dt}$$
 ,  $F_x = \iint_D K_x(x,y) \, dxdy$ ,  $F_y = \iint_D K_y(x,y) \, dxdy$ .

Соотношение (3) и есть искомое уравнение связи параметров движения изображений с их измеримыми характеристиками. Функционалы  $F_x$  и  $F_y$  называются индуцированными функционалами.

Дискретный вариант соотношения (3), использующий покадровое представление последовательности изображений, может быть записан как

$$F(E((N+1)T)) - F(NT) =$$

$$= \widetilde{F_{x}}(E(NT)) \, \widetilde{s_{x}} + \, \widetilde{F_{y}}(E(NT)) \, \widetilde{s_{y}}, \tag{4}$$

где T – кадровый период; N – номер кадра;  $\widetilde{s_x}$ ,  $\widetilde{s_y}$  – оценки смещений изображения по координатам;  $\widetilde{F_x}$ , .  $\widetilde{F_y}$  – усредненные по кадрам N и (N+1) значения соответствующих индуцированных функционалов. Далее будем называть кадр с номером N первым, а с номером (N+1) – вторым.

Система уравнений, разрешимая относительно компонент оптического потока, строится параметризацией соотношения (4). Одним из вариантов параметризации является изменение значения координат положения на изображениях областей определения функционала F. При таком подходе целесообразно, для построения требуемой по условиям разрешимости системы уравнений, назначить на анализируемых изображениях идентичные по форме и размерам подобласти анализа: А1 на первом, А2 - на втором кадре; и каждую из этих областей покрыть идентичной системой окон анализа: Dk<sub>i</sub>; k =1, 2 — номер подобласти, на которой организовано покрытие; i = 1, 2, ..., m – номер окна анализа. Для каждой пары окон анализа D1<sub>i</sub> и D2<sub>i</sub> с идентичным номером і строится уравнение вида (4). В результате, в общем случае получается переопределенная система уравнений, решение которой дает искомые оценки  $\widetilde{s_x}$ ,  $\widetilde{s_v}$  относительного смещения подобластей А1 и А2 в кадрах. Методика решения указанной системы уравнений и итерационная процедура высокоточного совмещения изображений приведена в [15]. В результате такого пофрагментного совмещения изображений формируется векторное поле скорости движения фрагментов изображе-

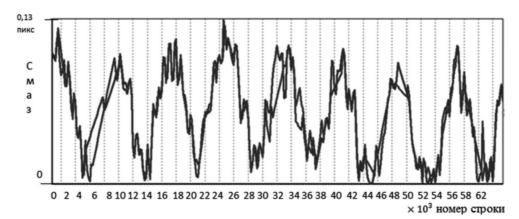


Рис. 2. Пример колебаний изображения, проецируемого на фокальную плоскость КА «Ресурс-П»

ния — оптический поток. Оптический поток характеризуется как параметрами векторов скорости движения фрагментов изображения, так и плотностью, то есть количеством вычисленных векторов на единицу площади изображения. Метод функционализации позволяет варьированием размерами области определения D функционала вычислять оптический поток с предельно возможной плотностью.

При контроле качества получаемых изображений интерес представляет не сама скорость движения изображения, имевшая место при съемке, а смаз видеоданных, получившийся за счет отклонения этой скорости от программно задаваемого значения. Поэтому ординаты всех графиков, приведенных ниже, градуированы в пикселях значения смаза.

На рис. 2 представлены графики колебаний изображения подстилающей поверхности на маршруте съемки КА типа «Аист-2». Полученные графики смаза на маршруте съемки показывают, что изображение в фокальной плоскости движется с колебаниями, вызванными внешними возмущающими воздействиями, к которым относятся воздействия со стороны системы управления движением КА, колебания солнечных батарей, высокочастотные вибрации устройств с вращающимися роторами. Амплитуда колебаний калибрована в пикселях смаза получаемого изображения, возникающего из-за колебаний изображения относительно заданного закона сканирования. По оси абсцисс – время, калиброванное в строках изображения, получаемого при сканирующей съемке. Частота получения данных измерений – 1843 Гц.

На рис. 2: ордината — это пиксели смаза изображения; абсцисса — это время, калиброванное в строках изображения, получаемого при сканирующей съемке. Различные графики соответствуют колебаниям изображения на краях различных фоточувствительных матриц, составляющих ОЭП.

Частотный спектр колебаний изображения на фокальной плоскости (рис. 2) представленных на рис. 3.

Проведенные экспертные оценки показывают: низкочастотные колебания 0,11 Гц вызваны процессом входа КА на маршрут съемки; частота 0,56 Гц обусловлена отработкой программы наведения системой управления КА; частота 11,5 Гц соответствует собственной частоте колебаний солнечных батарей. Частота 159,8 Гц вероятнее всего вызвана дисбалансом вращающихся роторов маховиков системы управления движением. Как

видно из приведенных графиков, вибрации изображения в фокальной плоскости приводят к появлению весьма незначительного смаза изображения.

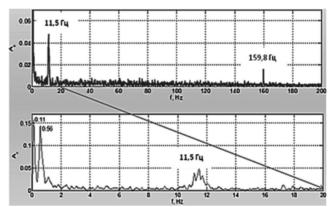


Рис. 3. Амплитудная частотная характеристика колебания изображения

В [18] предложена методика восстановления угловых скоростей движения оптической оси КА на основе информации о поле скоростей движения изображения (оптическом потоке) и положения центра масс КА на маршруте съемки. Полученное уточнённое значение угловой скорости движения оптической оси ОЭТК на маршруте съемки перспективно применять в задаче уточнения геодезической привязки космических изображений.

#### Заключение

Представлена комплексная технология диагностики состояния бортовой измерительной и целевой съемочной аппаратуры систем Д33.

Рассмотрено два аспекта при реализации такой технологии. Первый аспект связан с диагностированием и уточнением элементов внутреннего и внешнего ориентирования бортовой аппаратуры и КА в целом на основе анализа данных по наземным опорным полигонам. Второй аспект основан на прецизионном анализе параметров движения изображения в фокальной плоскости телескопа с использованием метода функционализации. Этот метод позволил определить частотный спектра колебаний фокальной плоскости целевой аппаратуры КА, вызванных внешними и внутренними возмущениями.

Представлены статистически достоверные данные по точности технологии комплексной диагностики с использованием натурной информации, полученной в

ходе эксплуатации космических систем Д33 «Ресурс-П» и «Аист-2».

### Литература

- 1. Кровотынцев В.А, Успенский С.А., Семериков А.Н., Леонов М.С. Наземный комплекс приема, обработки и распространения информации «Конопус-В» [Электронный ресурс] jurnal.vniiem.ru/text/148/24-35.pdf, 2015.
- 2. Ромашкин В.В. Комплексы приема, обработки, распределения и доведения до потребителей информации дистанционного зондирования Земли. Разработки ОАО «НИИ ТП» для космического комплекса «Ресурс-П» [Электронный ресурс] http://www.niitp.ru/component/content/article/56-dzz, 2017.
- 3. Абрамов Н.С., Талалаев А.А., Фраленко В.П., Хачумов В.М., Шишкин О.Г. Высокопроизводительная нейро-сетевая система мониторинга состояния и поведения подсистем космических аппаратов по телеметрическим данным [Электронный ресурс] http://psta.psiras.ru/read/psta2017\_3\_109-131.pdf, 2017.
- 4. Ахметов Р.Н. Управление живучестью низкоорбитальных автоматических КА ДЗЗ. // Аэрокосмический курьер. 2010. №6. С.2-4.
- 5. Ахметов Р.Н., Макаров В.П., Соллогуб А.В. Концепция автономного управления живучестью автоматических космических аппаратов дистанционного зондирования Земли в аномальных ситуациях. // Механика и машиностроение. Известия Самарского научного центра Российской академии наук. 2009. № 3. С.165-176.
- 6. Grodecki J. and Dail G. IKONOS Geometric Accuracy [Электронный pecypc] https://legacy.satimagingcorp.com/media/pdf/IKONOSGeometricAccuracyValidation-ISPRS202 002.pdf, 2002.
- 7. Kwoh L.K., Tan W.J. International Archives of the Photogrammetry, Remote Sensing and Spatial Information, Development of Camera Model and Geometric Calibration/Validation of XSAT IRIS Imagery. 2012. Vol. XXXIX-B1. pp. 239-243.
- 8. Malthus T., Fuqin Li. Calibration of Optical Satellite and Airborne Sensors [Электронный ресурс]. http://www.science.org.au/reports/index, 2014.
- 9. Современные технологии обработки данных дистанционного зондирования Земли / Под ред. Еремеева В.В. М.: Физматлит, 2015. 460 с.
- 10. Ахметов Р.Н., Еремеев В.В., Кузнецов А.Е., Мятов Г.Н., Пошехонов В.И., Стратилатов Н.Р. Высокоточ-

- ная геодезическая привязка изображений земной поверхности от КА «Ресурс-П» // Исследование Земли из космоса. 2017. № 1. С. 44-53.
- 11. Кузнецов А.Е., Пресняков О.А., Мятов Г.Н. Структурное восстановление изображений от многоматричных сканерных датчиков ДЗЗ // Цифровая обработка сигналов. № 3. 2015. С. 15-22.
- 12. Ахметов Р.Н., Зинина И.И., Юдаков А.А., Еремеев В.В., Кузнецов А.Е., Пошехонов В.И., Пресняков О.А., Светелкин П.Н. Точностные характеристики выходной продукции высокого разрешения КА «Ресурс-П». // Современные проблемы дистанционного зондирования Земли из космоса. № 3 Т.17. 2020. С. 142-148.
- 13. Кузнецов П.К., Мартемьянов Б.В., Мятов Г.Н., Юдаков А.А. Методика вычисления оценок параметров смаза изображений, получаемых целевой аппаратурой КА типа «Ресурс» // Тез. докл. междунар. научн.-техн. конф. «Космонавтика. Радиоэлектроника. Геоинформатика». Рязань: РГРТУ. 2017. С. 289-290.
- 14. Кузнецов П.К., Мартемьянов Б.В., Ращупкин А.В. Техническое зрение подвижных объектов. Методика совмещения изображений, полученных при наблюдении с подвижных объектов // Вестник компьютерных и информационных технологий. 2014. № 3. С. 3-10.
- 15. Кузнецов П.К., Мартемьянов Б.В., Семавин В.И. Техническое зрение подвижных объектов. Метод анализа поля скоростей динамического изображения // Вестник компьютерных и информационных технологий. 2014. № 1-М. С. 3-9.
- 16. Кузнецов П.К., Мартемьянов Б.В. Математическая модель формирования видеоданных, получаемых с использованием сканирующей съемки // Известия Самарского научного центра Российской академии наук. 2014. №6. С. 292-299.
- 17. Владимиров В.С. Уравнения математической физики. изд. 4-е. М.: Наука. Главная редакция физикоматематической литературы. 1981. 512 с.
- 18. Сомов Е.И., Бутырин С.А., Кузнецов П.К., Мартемьянов Б.В. Методика уточнения угловой ориентации автономной мобильной платформы на основе комплексирования измерительной и наблюдательной информации, не содержащей реперов // 5 российская мультиконференция по проблемам управления. Материалы конференции «Управление в технических, эргатических, организационных и сетевых системах». Санкт-Петербург: Концерн «ЦНИИ Электроприбор». 2012. С. 810-813.

14

УЛК 004.932

# КЛАСТЕРИЗАЦИЯ ГИПЕРСПЕКТРАЛЬНЫХ СПУТНИКОВЫХ СНИМКОВ ЗЕМНОЙ ПОВЕРХНОСТИ НА ОСНОВЕ МЕТОЛА ПЛОТНОСТИ БЛИЖАЙШИХ СОСЕЛЕЙ

Ларионов С.М., м.н.с. НИИ «Фотон» РГРТУ, e-mail: foton@rsreu.ru; Макаренков А.А., к.т.н., с.н.с. НИИ «Фотон» РГРТУ, e-mail: foton@rsreu.ru.

# CLUSTERING OF HYPERSPECTRAL SATILLITE IMAGES OF THE EARTH'S SURFACE BASED ON THE NEAREST NEIGHBORS DENSITY METHOD

## Larionov S.M., Makarenkov A.A.

The paper describes modified approach of clustering of Earth remote sensing hyperspectral data based on nearest neighbors density method. Features of this approach, its advantages and limitations are considered. The results of experimental studies are presented.

**Key words:** Earth remote sensing, hyperspectral images, image clustering, nearest neighbors method, nearest neighbors density.

**Ключевые слова:** дистанционное зондирование Земли, гиперспектральные снимки, кластеризация изображений, метод ближайших соседей, плотность ближайших соседей.

Описывается модификация подхода к кластеризации гиперспектральных данных дистанционного зондирования Земли на основе метода плотности ближайших соседей. Рассматриваются особенности данного подхода, его преимущества и ограничения. Приводятся результаты экспериментальных исследований.

#### Введение

На сегодняшний день дистанционное зондирование земной поверхности из космоса позволяет решать ряд важных тематических задач в картографии, метеорологии, изучении растительности и водных ресурсов, исследовании состава атмосферы и почвы, добыче полезных ископаемых, сельском хозяйстве, предупреждении и ликвидации последствий чрезвычайных ситуаций, наблюдении за антропогенными объектами и в других отраслях экономики и научных областях [1-6]. В последнее время на отечественных и зарубежных космических аппаратах устанавливается гиперспектральная аппаратура [7]. Каждая точка гиперспектрального изображения представляет собой спектральную характеристику, определяющую физико-химический состав малого участка подстилающей поверхности [8], что позволяет решать задачу классификации и оценки параметров объектов сцены с более высокой точностью по сравнению с обработкой традиционной мультиспектральной и панхроматической информации.

#### Подходы к классификации изображений

Методы классификации изображений можно разделить на две группы: на обучаемые и необучаемые подходы [8, 9]. Обучаемые подходы, например, такие как нейросети [10, 11], требуют значительного объема дополнительной информации в виде обучающих выборок. Подготовка «обучающих» изображений является сложным, время затратным процессом, не имеющим единственного решения. Результаты такой разметки во многом зависят от навыков и интуитивных предположений эксперта о принадлежности каждой конкретной точки изображения к той или иной смысловой группе. Таким образом получение обучающей выборки обычно требу-

ет существенных финансовых затрат [12]. При этом эффективность обучаемых методов кластеризации напрямую зависит от объема и качества обучающей выборки [8]. Поэтому большое внимание со стороны научного сообщества обращено в сторону необучаемых методов классификации, то есть кластеризации изображений. Под кластеризацией [13] понимается разбиение изображения на хорошо разделимые в выбранном пространстве признаков смысловые области точек, обладающих некоторой степенью общности. Известен ряд методов кластеризации изображений, некоторые из которых рассмотрены далее.

**Метод к-средних.** На первом шаге метода [14] произвольным образом выбираются начальные центры кластеров. Затем каждая точка изображения относится к тому кластеру, центр которого находится к ней (точке) ближе всего в пространстве признаков. После чего центры кластеров пересчитываются как центры масс кластеров. Эта процедура переразметки точек и перерасчета центров кластеров повторяется до тех пор, пока положение центров кластеров не стабилизируется. Одна из проблем данного метода в том, что существует вероятность схождения к локальному, а не глобальному оптимуму. К тому же необходимы априорные данные о числе классов на изображении, а результаты кластеризации зависят от начального выбора центров кластеров.

Метод нечеткой кластеризации с-средних. Согласно данному подходу [15] вначале задаются центры кластеров. Затем рассчитывается расстояние в выбранном пространстве признаков от каждой точки до центра каждого кластера и степень принадлежности каждой точки каждому из кластеров. На основе этих значений вычисляется некоторая целевая функция. За искомое принимается разбиение, обеспечивающие минимум вы-

бранной целевой функции. Основной проблемой метода является сильная зависимость результатов от выбора первоначальных центров кластеров, а также существенные вычислительные затраты.

**Метод DBSCAN.** Согласно данному методу [16], вначале необходимо задать некоторую *е*-окрестность (расстояние *е*) в выбранном пространстве признаков, в пределах которой определяются точки кластера, а также задать минимальное число *mPtrs* точек, образующих кластер. Если в *е*-окрестности некоторой точки находится как минимум *mPtrs точек*, то анализируемая точка считается *плотной*.

Алгоритм состоит из следующих шагов.

*Шаа 1*. Выбирается непомеченная точка изображения.

*Шаа 2.1.* Находятся все точки, лежащие в e-окрестности выбранной точки.

*Шаг 2.2* Если количество найденных точек меньше *mPtrs*, то данная точка помечается как *шум* и осуществляется переход к *Шагу 1*.

*Шаг 2.3.* В противном случае (то есть когда данная точка является *плотной*) данная точка и точки, лежащие в её e-окрестности, объединяются в кластер.

*Шае 3.1.* Среди найденных на *Шаее 2.1* точек, лежащих в e-окрестности рассматриваемой точки, также определяются *плотные* точки.

*Шаг 3.2.* Точки, лежащие в *е*-окрестности каждой *плотной* точки, найденной на *Шаге 3.1*, добавляются в данный кластер.

Шаг 3.3 Процедура Шаг 2.1-Шаг 3.2 повторяется для каждой из добавленных точек до тех пор, пока среди них не останется плотных. Это означает, что кластер сформирован и необходимо перейти к следующей неразмеченной точке изображения, повторив Шаги 1-3.3.

Минусом подхода является невозможность корректной кластеризации при условии разной плотности реально присутствующих на изображении классов объектов, так как в данном случае сложно подобрать подходящие e-окрестность и число mPtrs.

**Метод OPTICS** [17] является развитием метода DBSCAN. На первом шаге также задается расстояние e в выбранном пространстве признаков и число mPtrs. Для каждой точки v находится основное расстояние  $e_c(v)$  — наименьшее расстояние, при котором анализируемая точка v является nnomhoŭ, причем рассматриваются только  $e_c(v) \le e$ . Если такого  $e_c(v)$  не находится, то v помечается как mym. Затем для каждой точки определяется расстояние достижимости  $e_r(v)$  в соответствии с выражением:

$$e_r(v) = \begin{cases} \text{не определено, если } w\text{- помечена как } \textit{шум} \\ \max[e_c(w), \textit{dist}(v, w)], \text{ иначе,} \end{cases}$$
 (1)

где  $\max[\bullet]$  – оператор определения максимального значения; dist(v,w) – расстояние между точками v и w.

Согласно *OPTICS* создается упорядоченная база данных, в которой для каждой точки сохраняются основное расстояние и расстояние достижимости.

Такой подход позволяет избежать проблемы, связанной с различной дисперсией физических классов

изображения и выделить все классы, которые занимают в выбранном пространстве признаков окрестность, меньше или равную заранее заданному e. Однако этот подход требует больших вычислительных затрат, в особенности применительно к гиперспектарльным изображениям, содержащим миллионы многомерных элементов.

**ЕМ-алгоритм.** Данный подход [13] требует априорного знания функции плотности вероятности для каждого физического класса с соответствующим значением математического ожидания и дисперсией. Результаты кластеризации сильно зависят от первоначальной оценки вида распределения.

В публикациях [18-20] рассматривается подход к кластеризации гиперспектральных данных на основе вычисления плотности ближайших соседей в выбранном пространстве признаков (Graph WatershEd using Nearest Neighbors — метод GWENN). Главное преимущество этой группы алгоритмов в том, что для их применения не требуется каких-либо предварительных знаний о структуре данных, таких как фактическое количество физических классов на изображении, статистическое распределение исходных данных, экземпляры классов или обучающие выборки.

#### Метод на основе плотности ближайших соседей

Идея подхода (метод GWENN) [18-20] заключается в определении метки каждой точки изображения исходя из меток ближайших соседей и плотностей этих ближайших соседей в выбранном пространстве признаков. На этапе предобработки выбирается число k ближайших соседей, пространство признаков, мера расстояния  $\Delta$  и функция плотности точки  $\omega$ .

В качестве пространства признаков могут быть выбраны значения спектральной плотности энергетической яркости (СПЭЯ) точек гиперспектрального изображения [8].

За меру расстояния  $\Delta$  может приниматься, например, Евклидова метрика:

$$\Delta(x, y; i, j) = \sqrt{\sum_{h}^{H} (I_{h}(x, y) - I_{h}(i, j))^{2}},$$
 (2)

где (x,y) , (i,j) — координаты точек на изображении I;  $\Delta(x,y;i,j)$  — расстояние между точкой (x,y) и точкой (i,j) ;  $I_h(x,y)$  ,  $I_h(i,j)$  — значения СПЭЯ в h-м канале точки (x,y) и точки (i,j) .

Плотность  $\varphi(x,y)$  может быть выбрана как величина, обратно пропорциональная суммарному расстоянию от точки (x,y) до каждой из её k ближайших соседей:

$$\varphi(x,y) = \frac{1}{\sum_{(i,j) \in RNN(x,y)} \Delta(x,y;i,j)},$$
(3)

где kNN(x,y) — множество k ближайших соседей точки (x,y) .

В ряде источников функция  $\phi(x,y)$  выбирается таким образом, чтобы вес каждого слагаемого в знаменателе выражения (3) убывал в зависимости от «удаленности» в пространстве признаков соответствующего «соседа» от данной точки.

Согласно методу GWENN, вначале для каждой точки изображения вычисляется массив расстояний от анализируемой точки до каждой другой точки изображения. Элементы массива сортируются в порядке возрастания. Для каждой точки с координатами (x, y) определяются её k ближайших соседей и её плотность  $\varphi(x,y)$ . Затем точки ранжируются в порядке убывания значения плотности. Предполагается, что наиболее плотные точки являются типичными экземплярами физических классов, поэтому процесс кластеризации начинается именно с них (они являются первыми элементами формируемых кластеров). Принадлежность к какому-либо физическому классу точек, менее плотно окруженных своими соседями в пространстве признаков, определить сложнее, поэтому их кластеризация производится, в последнюю очередь, когда большинство их ближайших соседей уже имеют метки. Также точки с наименьшей плотностью могут быть помечены как шумовые.

Этап кластеризации точек, согласно *методу GWENN*, состоит из следующих шагов.

*Шаа* 1. Выбирается наиболее плотная точка (x,y), не имеющая метки.

Шаг 2.1. Если среди множества kNN(x,y) нет точек, имеющих метку, точке (x,y) присваивается новая метка (число кластеров увеличивается на единицу) и осуществляется переход к Шагу 1. Метки принимают целые положительные значения, начиная с 1. Метка каждого нового кластера равняется текущему числу кластеров, увеличенному на единицу.

Шаг 2.2. Если среди множества kNN(x,y) есть точки, имеющие метку, то точке (x,y) присваивается метка, соответствующая наибольшей суммарной плотности среди kNN(x,y):

$$Lbl(x,y) = \arg\max_{l \in [1,L]} \left[ \sum_{\substack{(i,j) \in kNN(x,y) \\ Lbl(i,j) = l}} \varphi(i,j) \right], \tag{4}$$

где Lbl(x,y) — метка точки (x,y); l — значение метки; L — текущее число кластеров;  $\varphi(i,j)$  — плотность точки (i,j);  $\sum_{\stackrel{(i,j)\in kNN(x,y)}{lbl(i,j)=l}} \varphi(i,j)$  — сумма плотностей точек, имею-

щих метку l, из множества kNN(x,y) ближайших соседей точки (x,y). После того, как точка (x,y) получила метку, осуществляется переход к  $extit{\it Wasy 1}$ .

Основным недостатком подхода является неоднозначность выбора числа k, непосредственно влияющего на точность кластеризации и на требуемые вычислительные затраты. С одной стороны, чтобы добиться высокого быстродействия алгоритма, k должен принимать как можно меньшее значение. Однако, например, проведенные экспериментальные исследования на изображении, полученном датчиком AVIRIS над долиной Салинас (изображение *corrected Salinas*), размером 512 строк на 217 столбцов и включающим 16 физических классов объектов, показали следующие результаты. При выборе k, равном 0,072 % от общего числа то-

чек, в результате кластеризации получены 214 кластеров. То есть физические классы, реально присутствующие на изображении, оказались разбиты на большое число малых кластеров, что в свою очередь затрудняет идентификацию объектов. Для повышения точности кластеризации, необходимо увеличить значение к. Однако при этом растет время обработки (или требуемая оперативная память), к тому же кластеры становятся более крупными и включают в себя уже несколько физических классов объектов. Например, для изображения corrected Salinas при выборе k, равным 3.6 % от общего числа точек получено 7 кластеров. То есть более мелкие физические классы оказались объединены с более крупными, что недопустимо. Рассмотренный метод при кластеризации натурной информации (маршрутов размером 1000х3000 пикселей) предъявляет высокие требования к аппаратным средствам (при оптимизации по времени выполнения требуемый объем оперативной памяти составляет ~43 Гб).

Известна реализация описываемого алгоритма [19] на основе пирамидальной обработки, состоящая из следующих шагов.

Шаа 1. Получение пирамиды изображений. Каждое изображение s-го уровня формируется путем вейвлет-преобразования изображения предыдущего (s-1)-го уровня. Изображение s-го уровня является более грубым по сравнению с изображением (s-1)-го уровня. Каждой точке s-го уровня соответствует 4 точки изображения (s-1)-го уровня.

Шаг 2. Кластеризация нижнего уровня пирамиды при помощи метода GWENN. Выделение наиболее плотной точки для каждого полученного кластера — т.е. нахождение экземпляров кластеров.

Шае 3.1. Кластеризация изображения следующего s-го уровня, в ходе которой при нахождении ближайших соседей используются точки изображения (s-1)-го уровня, соответствующие экземплярам кластеров изображения s-го уровня.

*Шаг* 3.2. Выделение экземпляров кластеров на изображении s-го уровня. Повторение процедуры *Шаг* 3.1 — *Шаг* 3.2 до уровня исходного изображения.

Пирамидальная реализация требует меньших вычислительных затрат, однако приводит к повышенной, по сравнению с непирамидальным методом GWENN, потере точек малых классов.

Для решения описанных выше проблем предлагается модификация *метода GWENN*, состоящая из следующих шагов.

Шае 1. Кластеризация исходного изображения с малым k (в экспериментах использовалось k = 0,01 % от числа N точек изображения). В результате получается большое число первичных кластеров, каждый из которых содержит малое количество элементов. Физические классы объектов на изображении оказываются разделены на несколько первичных кластеров.

*Шаг* 2. Нахождение «средних спектров» для каждого первичного кластера:

$$\overline{I_h}(l) = \frac{1}{N_l} \sum_{(x,y)} I_h^l(x,y), \quad l = \overline{1,L},$$
 (5)



а) Исходное изображение



в) Результат кластеризации методом GWENN



 г) Результат кластеризации пирамидальным (двухуровневым) GWENN



б) Эталон кластеризации



д) Результат кластеризации
 предлагаемой модификацией GWENN

Рис. 1. Кластеризация изображения corrected Salinas

где  $\overline{I_h}(l)$  – среднее значение СПЭЯ точек l-го первичного кластера в h-м канале;  $I_h^l(x,y)$  – значение СПЭЯ в h-м канале точки с координатами (x,y), имеющей первичную метку l;  $N_l$  – число точек, имеющих первичную метку l; L – число первичных кластеров.

*Шав* 3. Определение оптимального  $k'_{opt}$  в смысле выбранной меры разделимости R(k',way) для объединения «средних спектров»:

$$k'_{opt} = \max_{\substack{k' \in [1,L] \ \text{operator}}} (R(k', way)),$$
 (6)

где k' — число ближайших соседей на этапе объединения первичных кластеров; way — параметр, определяющий направление кластеризации: ascend — от «менее плотных» точек к «более плотным», descend — от «более плотных» точек к «менее плотным».

*Шае 4*. Кластеризация «средних спектров» с числом  $k_{out}'$  ближайших соседей.

*Шаг* 5. Получение итоговых кластеров изображения — переразметки малых первичных кластеров в соответствии с результатами кластеризации их «средних спектров».

Для оптимизации числа k' ближайших соседей находится максимум функции следующего вида:

$$R(k',way) = \frac{\sigma_{\text{inter}}^2(k',way)}{\sigma_{\text{intra}}^2(k',way) \cdot NC'(k',way)},$$
(7)

где  $\sigma_{\text{inter}}^2(k',way)$  — среднее расстояние между кластерами в выбранном пространстве признаков;  $\sigma_{\text{intra}}^2(k',way)$  — среднее расстояние внутри кластеров в выбранном пространстве признаков; NC'(k',way) — число полученных кластеров; t — показатель степени числа кластеров.

Мера разделимости R(k',way) учитывает внутрикластерное и межкластерное расстояние и число кластеров и обеспечивает компактность и отделимость результирующих кластеров в выбранном пространстве признаков. При кластеризации «средних спектров» использу-

ется проход как от «среднего спектра» с наибольшей плотностью, так и в обратном направлении (параметр way). Предполагается, что «средний спектр», соответствующий некоторому первичному кластеру и отстоящий от других «средних спектров» в пространстве признаков, не является шумовым, а соответствует малому классу объектов на изображении. Данный подход позволяет назначить такому малому физическому классу уникальную метку и таким образом сохранить его в итоговых кластерах, не объединяя с более крупными. Для снижения вычислительных затрат можно при кластеризации «средних спектров» перебирать не все значения  $k' \in [1, L]$ . Если при некотором значении k' все «средние спектры» объединяются в один кластер (то есть все точки изображения относятся к одному итоговому кластеру), дальнейшее увеличения значения k' не имеет смысла, потому как при этом «средние спектры» так же будут объединяться в один кластер.

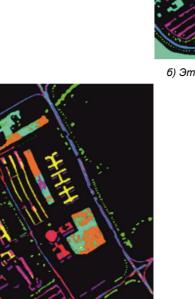
В табл. 1 и на рис. 1 и рис. 2 представлены результаты экспериментальных исследований.

#### Заключение

Проведенные экспериментальные исследования предлагаемой модификации метода GWENN показали повышение точности кластеризации и уменьшение затрат памяти. Применение предлагаемого алгоритма позволяет корректно кластеризовать присутствующие классы объектов, занимающие малую часть сцены (таким образом избежать потери информации). Дальнейшие исследования авторов направлены на подбор меры разделимости R, обеспечивающей необходимые точностные параметры и низкие вычислительные затраты при работе с гиперспектральными изображениями с различными соотношением размеров объектов наблюдаемой сцены и степенью однородности их спектральных характе-



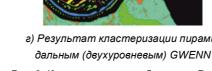
а) Исходное изображение



г) Результат кластеризации пирами-



д) Результат кластеризации предлагаемой модификацией GWENN



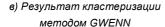


Рис. 2. Кластеризация изображения Pavia University

. Результаты экспериментальных	

	Memod GWENN	Пирамидальный (двухуровневый) <i>GWENN</i>	Предлагаемая модификация GWENN
Изображение correcte	ed Salinas (512 строк, 21	7 столбцов, 214 каналов, 16 физі	ических классов) [21]
Доля верно кластеризованных	70,1251 %	67,7142 %	72,4750 %
точек	(k = 1200)	(k = 290)	(t = 0.73)
Требуемый объем памяти	1,6 GB	98 MB	66 MB
Время обработки	8 мин. 53 сек.	41 сек.	10 мин. 16 сек.
Изображение Pavia U	University (610 строк, 34	0 столбцов, 103 каналов, 13 физи	ческих классов) [21]
Доля верно кластеризованных	65,43 %	63,17 %	68,01 %
точек	(k = 1500)	(k = 350)	(t = 0.87)
Требуемый объем памяти	3 GB	178 MB	102 MB
Время обработки	15 мин. 8 сек.	1 мин. 38 сек.	18 мин. 27 сек.

Примечание. Экспериментальные исследования проводились на персональном компьютере со следующими характеристиками:

- процессор Intel Core i7, 3.6 ГГц;
- объем оперативного запоминающего устройства 32 Гб
- операционная система Windows 7 Профессиональная (64-разрядная).

#### Литература

- 1. Загородная А.А. Опыт применения материалов ДЗЗ в работах экологической тематики / Н.Б. Пыстина, А.А. Загородная // Защита окружающей среды в нефтегазовым комплексе. 2005. № 9. С. 42-48.
- 2. Сай С.В. Алгоритмы морфологической кластеризации растительности на базе космических снимков применительно к территории г. Хабаровска / С.В. Сай, Г.Я. Маркелов, С.В. Плесовских // Вестник Тихоокеанского госурадственного университета. 2017 № 3 (46). С. 13-22.
- 3. Клейн Н.А. Анализ подтопленных сельскохозяйственных территорий в Западной Сибири с использованием ДЗЗ / Н.А. Клейн, Л.В. Березин, М.С. Балуков // Инновации в природообустройстве и защите в чрезвычайных ситуациях: Материалы IV международной научно-практической конференции. 2018. С. 88-91.
- 4. Кениг А.В. Применение технологии ГИС и методов ДЗЗ в системе выявления, мониторинга и охраны объектов археологического наследия / А.В. Кениг, Е.А. Зайцева // ИНТЕРЭКСПО ГЕО-Сибирь. 2013. Т. 8. С. 48-51.
- 5. Уфимцев А.Е. Изучение структуры природных и антропогенных ландшафтов с применением ГИС и данных ДЗЗ / А.Е. Уфимцев, О.Ю. Вавер // Бассейновые территории: проблемы и пути их решения: Материалы международной научно-практической конференции. редактор-составитель Г.С. Кощеева. 2013. 2016. С. 56-60.
- 6. Абросимов А.В. Перспективы применения данных ДЗЗ из космоса для повышения эффективности сельского хозяйства в России / А.В. Абросимов, Б.А. Дворкин // Геоматика. 2009. № 4. С. 46-49.
- 7. Третьяков В.А., Основные тенденции развития гиперспектральной аппаратуры в мире / В.А. Третьяков // Космонавтика и Ракетостроение. 2013. № 4 (73). С. 36-40.
- 8. Advances in Hyperspectral Image and Signal Processing: A Comprehensive Overview of the State of the Art. / P. Ghamisi, N. Yokoya, J. Li, W. Liao, S. Liu, J. Plaza, B. Rasti, A. Plaza // IEEE Geoscience and Remote Sensing Magazine. 2017. December. Vol. 5, no. 4 Pp. 37-78. DOI:10.1109/mgrs.2017.2762087.
- 9. Hyperspectral Remote Sensing Classifications: A Pers-pective Survey / D. Chutia, D.K. Bhattacharyya, K.K. Sarma, R. Kalita, S. Sudhakar // Transactions in GIS. 2016. August. Vol. 20, no. 4. Pp. 463-490.
  - 10. Zhang S.L. A Study of Image Classification of Remote Sens-

- ing Based on Back-Propagation Neural Network with Extended Delta Bar Delta / S.L. Zhang, T.C. Chang // Mat-hematical Problems in Engineering. 2015. October. DOI: 10.1155/2015/178598.
- 11. Convolutional Neural Networks Based Hyperspectral Image Classification Method with Adaptive Kernels / C. Ding, Y. Li, Y. Xia, W. Wei, L. Zhang, Y. Zhang // 2017. June. Vol. 9, no. 6.
- 12. Towards a completely blind classifier for hyperspectral images / P. Halle; S. Le Moan, C. Cariou // International Conference on Image and Vision Computing New Zealand (IVCNZ). 2017.
- 13. Черезов Д.С. Обзор основных методов классификации и кластеризации данных / Д.С.Черезов, Н.А. Тюкачев // Вестник ВГУ, серия: системный анализ и информационные технологии. 2009. № 2. С. 25-29.
- 14. Arthur D. How slow is the k-means method? / D. Art-hur, S. Vassilvitskii // Proceedings of the 2006 Symposium on Computational Geometry. 2006. Pp. 144-153.
- 15. Fuzzy c-means algorithm for segmentation of aerial photography data obtained using unmanned aerial vehicle / M.V. Akinin, N.V. Akinina, A.Y. Klochkov, M.B. Nikiforov, A.V. Sokolova // International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. 2015. Vol. XL-5/W6. Pp. 113-115.
- 16. DBSCAN revisited, revisited: why and how you should (still) use DBSCAN / E. Schubert, J. Sander, M. Ester, H.P. Kriegel, X. Xu // ACM transactions on database sys-tems. 2017. Vol. 42 Nº 3 Pp. 1-21.
- 17. OPTICS: ordering points to identify the clustering structure / M. Ankerst, M. M. Breunig, H.-P. Kriegel, J. Sander//Proc. 1999 ACM SIGMOD Intern. Conf. on Management of data. –1999. Pp. 49-60.
- 18. Cariou C. A new k-nearest neighbor density-based clustering method and its application to hyperspectral images / C. Cariou, K. Chehdi // International Geoscience and Remo-te Sensing Symposium (IGARSS). -2016.
- 19. Cariou C. Nearest neighbor-density-based clustering methods for large hyperspectral images / C. Cariou, K. Chehdi // Proc. SPIE 10427, Image and Signal Processing for Remote Sensing XXIII. 2017. October. DOI: 10.1117/12.2278221.
- 20. Cariou C. Application of unsupervised nearest neighbor density-based approaches to sequential dimensionality reduction and clustering of hyperspectral images / C. Cariou, K. Chehdi // Proc. SPIE 10789, Image and Signal Processing for Remote Sensing. 2018. October.
- http://www.ehu.eus/ccwintco/index.php?title=Hyper-spectral Remote Sensing Scenes.

УДК 004.932.2

# ПРОГРАММНОЕ УПРАВЛЕНИЕ МНОГОСЕРВЕРНОЙ ОБРАБОТКОЙ ИЗОБРАЖЕНИЙ БОЛЬШОГО ОБЪЕМА ОТ СИСТЕМ КОСМИЧЕСКОГО НАБЛЮДЕНИЯ ЗЕМЛИ

Кузнецов А.Е., д.т.н., профессор, зам. директора НИИ «Фотон» РГРТУ, e-mail: foton@rsreu.ru; Рыжиков А.С., м.н.с НИИ «Фотон» РГРТУ; e-mail: foton@rsreu.ru.

# MULTISERVER PROCESSING SOFTWARE MANAGEMENT FOR LARGE-SCALE REMOTE SENSING IMAGERY

# Kuznechov A.E., Ryzhikov A.S.

In recent years, the requirements for the efficiency of the remote sensing products formation have become more stringent. In this regard, the topical problem is the effective software organization of multi-server parallel processing of large-volume images from space observation systems of the Earth. The Hadoop technology is shown in the article to be inappropriate for solving the problem. A solution which provides an efficient organization of multi-server parallel processing of remote sensing data has been proposed.

Key words: distributed processing, remote sensing, Hadoop, MapReduce.

**Ключевые слова:** распределенная обработка, дистанционное зондирование Земли, Hadoop, MapReduce.

#### Введение

Общемировая тенденция развития систем ДЗЗ высокого пространственного разрешения связана с повышением детальности и оперативности получения информации о наблюдае-

мых территориях. Это, в свою очередь, приводит к увеличению объемов передаваемой на приемные центры информации и ужесточению требований к скорости ее обработки. Поэтому в настоящее время для получения выходных информационных продуктов стандартных уровней применяются серверные кластеры. Для сокращения временных затрат осуществляется распределение маршрутов съемки по серверам и их параллельная обработка. Такое решение оказывается весьма эффективным для маршрутов длительностью 10-15 секунд и объемом 15-20 Гбайт. Однако для маршрутов съемки длительностью порядка 100 секунд и более и объемом 200-300 Гбайт вычислительных ресурсов одного сервера оказывается недостаточно и временные затраты на получение выходной продукции становятся значительными. Выходом из этой ситуации является организация распределенной обработки фрагментов одного маршрута съемки на нескольких серверах. При этом добиться близкого к линейному снижения временных затрат оказывается сложно, поскольку требуется минимизировать накладные расходы, возникающие при межсерверном обмене данными, необходимыми для обработки каждого фрагмента исходного маршрута съемки. В связи с этим целью работы является анализ известных программных решений по организации многосерверной обработки, выбор и обоснование наиболее эффективного подхода.

#### Организация многосерверной обработки

Процесс обработки маршрута съемки, представленного многозональным изображением  $I_i(m,n)$  , где i –

В последние годы ужесточаются требования к оперативности формирования продуктов ДЗЗ. В связи с этим актуальна задача эффективной программной организации многосерверной параллельной обработки изображений большого объема от систем космического наблюдения Земли. В работе показано, что технология Надоор не подходит для решения поставленной задачи. Предложено решение, обеспечивающее эффективную организацию многосерверной параллельной обработки данных ДЗЗ.

номер спектральной зоны, и вектором входных параметров  $\overline{u}$ , связан с последовательным выполнением определенных технологических операций  $T_j$ , где j – идентификатор вида обработки [1]. Укрупненную модель получения изображения выходного информационного продукта  $D_i(x,y)$  представим в виде

$$D_{i}(x,y) = T_{\kappa}[T_{\varepsilon}[T_{omm}[T_{\upsilon}(I_{i}(m,n),\overline{u})]]],$$

где индекс  $\kappa$  обозначает операцию ортотрансформирования в картографическую проекцию, e — геодезическое ориентирование, omm — поиск опорных точек, p — радиометрическую коррекцию.

В этой модели операция  $T_p$  связана с поэлементным яркостным преобразованием исходного изображения  $I_i(m,n)$  и может выполняться как для всего маршрута съемки, так и для отдельных его фрагментов. Т.е. при ее выполнении достаточно распараллеливания вычислительного процесса.

Операция поиска опорных и одноименных точек  $T_{\mathit{отм}}$  предназначена для подготовки исходных данных для операции геодезического, межматричного и межзонального ориентирования [1]. Она так же может выполняться параллельно для каждого фрагмента исходного изображения  $I_i(m,n)$ .

Операция геодезического ориентирования  $T_{\scriptscriptstyle c}$  связана с уточнением по опорным и одноименным точкам элементов внешнего ориентирования, описываемых вектором  $\overline{u}$  . В результате выполнения этой операции гарантируется высокоточное совмещение спектрозональных

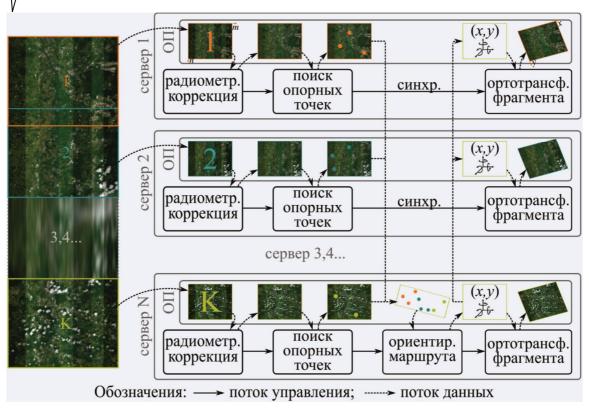


Рис. 1. Схема многосерверной обработки фрагментированного изображения

изображений  $D_i(x,y)$ . При выполнении этой операции для каждого фрагмента исходного изображения по сво-им исходным данным фрагменты выходного изображения  $D_i(x,y)$  не будут совмещаться между собой с субпиксельной точностью. Т.е. операция  $T_\varepsilon$  не распараллеливается.

Операция трансформирования в картографическую проекцию  $T_{\kappa}$  допускает независимую обработку каждого фрагмента по уточненным значениям вектора  $\overline{u}$  .

Схема распределенной обработки изображения маршрута съемки с учетом рассмотренных выше ограничений приведена на рис. 1. На этой схеме, в частности, показано, что геодезическое ориентирование маршрута съемки выполняется с привлечением опорных точек, найденных независимо на всех фрагментах, а параметры ориентирования далее распространяются на все сервера. Время выполнения геодезического ориентирования не зависит от протяженности маршрута, поэтому при организации многосерверной обработки по этой схеме достижимо ускорение обработки изображения, близкое к линейному. Для достижения максимального быстродействия основной задачей программной реализации является минимизация накладных расходов на межсерверное взаимодействие.

Программная реализация многосерверной обработки. В работах [2, 3] предлагается использовать Арасhе Hadoop реализацию парадигмы MapReduce в качестве основы для построения распределенной многосерверной обработки изображений. Анализ этих работ позволил выявить три особенности.

Первое. Входные и выходные данные, промежуточные результаты хранятся в файловой системе HDFS на накопителях. Это удобное свойство для задач, связан-

ных с обработкой сверхбольших (десятки ТБ) объемов слабо структурированных данных, на которые Наdоор изначально ориентирован. Однако объем спутниковых изображений высокого разрешения редко превышает 100 ГБ. Фрагменты таких изображений целесообразно обрабатывать с применением парадигмы резидентных вычислений, обращаясь только к данным в оперативной памяти, однако Нadoop не предоставляет такую возможность. Избыточные операции ввода/вывода в HDFS приводят к существенным временным затратам.

Второе. Наdoop предлагает API на языке Java, в то время как промышленная разработка современных вычислительно сложных приложений ведется на сравнительно низкоуровневых языках, таких как C, C++, Rust. Наличие сборщика мусора и виртуальной машины в Java создает труднопреодолимые ограничения для эффективной реализации вычислительных задач на этом языке. Интерфейс Hadoop Streaming позволяет интегрировать в Hadoop программы, написанные на любом языке. Но его использование связано с накладными расходами на конвертацию данных и их передачу между процессами.

Третье. Экосистема Hadoop содержит множество разнообразных абстракции и инструментов, ускоряющих разработку и отладку распределенных приложений. Однако почти все они ориентированы на обработку текстов. Готовых инструментов для загрузки, фрагментирования и доступа к данным изображений нет.

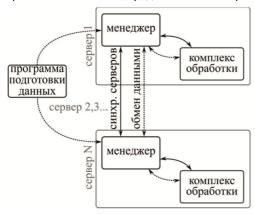
Для оценки степени влияния первых двух особенностей на производительность обработки на основе Hadoop Streaming было создано экспериментальное приложение, выполняющее загрузку изображения в HDFS, его разбиение на перекрывающиеся фрагменты, поворот каждого фрагмента на фиксированный угол и

объединение результатов в единый растр. Затраты вревремени на операции ввода-вывода составили около 60% всего времени работы, примерно 40% из которых пришлось на операцию объединения. При этом реализация объединения (reduce) нетривиальна, т.к. HDFS не поддерживает запись в произвольное место выходного файла.

В связи с выделенными недостатками для управления распределенной обработкой фрагментированного маршрута предлагается использовать специализированное программное решение на языке C++.

Для исключения избыточных операций вводавывода обеспечим на каждом сервере прямую запись входных фрагментов изображений в ОП минуя жесткий диск; хранение промежуточных данных в ОП. Для организации взаимодействия согласно схеме, приведенной на рис. 1, реализуем упрощенную модель MapReduce.

Структурно предлагаемое решение состоит из нескольких программных компонент, представленных на рис. 2.



Puc. 2. Схема взаимодействия программ при распределенной работе

Программа подготовки исходных данных выполняет распаковку и декомпозицию непрерывного изображения маршрута съемки на перекрывающиеся фрагменты. Эта программа помещает сформированные фрагменты в ФС Серh, хранящую данные в ОП кластера серверов. Менеджер обеспечивает управление работой комплекса обработки, межсерверный обмен данными, синхронизацию этапов обработки на разных серверах и оповещение о текущем прогрессе. Менеджер передает между серверами только необходимую для их совместной работы информацию — опорные точки, результаты ориентирования и служебную информацию, включающую квитанции синхронизации. Суммарно объем передаваемого менеджером трафика не превышает 20 Мб при объеме маршрута 40 Гб.

Комплекс обработки, реализованный на языке C++, работает под управлением менеджера и реализует алгоритмы обработки изображений маршрута для каждого этапа. Комплекс имеет прямой доступ к фрагменту изображения и промежуточным результатам, в т.ч. самостоятельно выполняет их аллокацию, что позволяет при необходимости задать требуемое выравнивание для векторных операций AVX(параллелизм на уровне данных), обеспечить DMA (прямой доступ к памяти) доступ при использовании графических ускорителей (CUDA) и т.п. без создания дополнительных копий.

#### Менеджер обработки

Для повышения быстродействия программа подготовки исходных данных выделяет из входного потока фрагменты на обработку, не дожидаясь его окончания. Поэтому в момент начала обработки первого фрагмента неизвестно, сколько всего фрагментов будет передано в обработку и сколько серверов будет задействовано. Эта информация появляется лишь после исчерпания входного потока и выдачи последнего фрагмента. В связи с этим менеджер обработки реализован в виде унифицированной программы, без выделения клиентской и серверной части и установлен на каждом сервере. Каждый менеджер анализирует сопровождающую информацию обрабатываемого фрагмента маршрута. Если фрагмент является последним, менеджер переходит в режим координатора. Все остальные менеджеры получают роль клиентов. Менеджеры-клиенты обеспечивают поэтапную обработку каждый своего фрагмента, передачу найденных опорных точек координатору и ожидание от него результатов ориентирования. Координатор также обрабатывает один фрагмент изображения (последний), но дополнительно реализует сбор опорных точек, выполнение этапа ориентирования и выдачу результатов менеджерам-клиентам.

На рис. 3 приведена диаграмма, иллюстрирующая последовательность выполнения этапов обработки. Синхронизация серверов выполняется только при ориентировании, что позволяет амортизировать возможную разницу во времени выполнения предшествующих этапов.

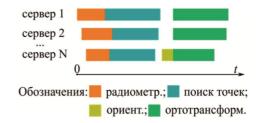


Рис. 3. Диаграмма последовательности выполнения этапов обработки

В реализованном на данный момент прототипе синхронизация реализована на основе файлов-квитанций в  $\Phi$ С Ceph. После завершения выполнения k-ого этапа менеджер-клиент помещает в разделяемую папку Ceph файл-квитанцию с уникальным именем, включающим имя этапа и номер обрабатываемого фрагмента. В случае, если при обработке возникла ошибка, в этот файл помещаются сведения об ошибке. Менеджер-координатор ожидает появления всех квитанций, свидетельствующих об успешном завершении выполнения к-го этапа, выполняет k+1-й и создает квитанцию, разрешающую клиентам запуск следующего этапа. Если хотя бы одна квитанция содержит информацию о критической ошибки, аналогичным способом координатор сигнализирует о прерывании обработки и возвращает код ошибки управляющей программе.

Синхронизация на основе квитанций просто реализуется, обеспечивает корректное совместное функционирование менеджеров обработки, но вероятно не является оптимальным решением с точки зрения быстродействия. В будущем планируется заменить файлы-кви-

танции на асинхронный опрос координатора посредством REST API.

#### Заключение

В предложенном программном решении отсутствует ввод-вывод с использованием дисковых накопителей. В отличие от Hadoop MapReduce, исходные фрагменты маршрута помещаются в ФС, построенную на базе ОП, после чего передается на сервер обработки и загружаются в ОП. Создаваемые далее в процессе обработки промежуточные изображения хранятся только в ОП сервера, обрабатывающего данный фрагмент и не передаются по сети на другие машины. Анализ результатов работы описанного ранее экспериментального приложения, реализованного в виде комплекса обработки в рамках предложенного решения, показал, что накладные расходы на межсерверное взаимодействие составили менее 7 % общего времени работы. При этом объем кодовой базы менеджера обработки на несколько порядков меньше объема исходного кода Hadoop, и как следствие, проще в поддержке и сопровождении.

Рассмотренное в статье программное решение по

организации многосерверной обработки маршрутов съемки большого объема является составной частью специального программного обеспечения обработки информации от спутников серии «Ресурс-П», «Ресурс-ПМ», РБКА и других систем ДЗЗ.

#### Литература

- 1. Р.Н. Ахметов, И.И. Зинина, А.А. Юдаков, В.В. Еремеев, А.Е. Кузнецов, В.И. Пошехонов, О.А. Пресняков, П.Н. Светелкин. Точностные характеристики выходной продукции высокого разрешения КА «Ресурс-П» // Современные проблемы дистанционного зондирования Земли из космоса. 2020. Т. 17. № 3. С. 41-47.
- 2. Almeer, M.H. Cloud Hadoop Map Reduce For Remote Sensing Image Analysis / M.H. Almeer // J. of Emerging Trends in Computing and Information Sciences. 2012. Vol. 3, no.4, pp. 637-644.
- 3. Wiley, K. Astronomical Image Processing with Hadoop / K. Wiley, A. Connolly, S. Krughoff, J. Gardner, M. Balazinska, B. Howe, Y. Kwon, Y. Bu // Astronomical Data Analysis Software and Systems XX. ASP Conference Proceedings. 2010. Vol. 442, pp. 93-98.

#### 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 — 250 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. Журнал «Цифровая обработка сигналов» включен в базу данных Web of Science - Russian Science Citation Index (287-я позиция)

### Планируемые сроки издания отдельных номеров журнала:

- № 4 декабрь 2020 г. Тематический выпуск: «ЦОС в радиотехнике и системах телекоммуникаций».
- № 1 март 2021 г. Тематический выпуск: ЦОС в инфокоммуникационных системах».
- № 2 июнь 2021 г. Тематический выпуск по материалам 23-й Международной научно-технической конференции «Цифровая обработка сигналов и ее применение-DSPA».
- № 3 сентябрь 2021 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

УЛК 004.932.4

# РАЗРАБОТКА НЕЙРОСЕТЕВЫХ МЕТОДОВ РАСПРОСТРАНЕНИЯ ЦВЕТА В ВИДЕО

Лаврушкин С.В., аспирант Московского государственного университета имени М.В. Ломоносова, e-mail: sergey.lavrushkin@graphics.cs.msu.ru

Ватолин Д.С., к.ф.-м.н., старший научный сотрудник Московского государственного университета имени М.В. Ломоносова, e-mail: dmitriy@graphics.cs.msu.ru

#### DEVELOPMENT OF NEURAL NETWORKS FOR COLOR PROPAGATION IN VIDEO

#### Lavrushkin S.V., Vatolin D.S.

The task of creating color versions of existing black and white images and videos originated immediately after color images and videos had emerged. The process of color prediction and propagation for grayscale images and videos is called colorization. While colorization has been gradually automated with the development of digital image and video processing, there is a lot of manual work. In this paper, we consider a problem of color propagation in video using a single colored frame, which can be obtained either by manual colorization or by applying automatic image colorization methods. To solve this problem we propose two types of neural network methods: the former are based on the prediction of separable filters, the latter use optical flow to propagate color between frames. The majority of proposed methods have shown high visual quality on test datasets with preservation of colorization quality for a large number of frames forward in the videos.

**Key words:** deep learning, convolutional neural networks, colorization.

**Ключевые слова:** глубокое обучение, сверточные нейронные сети, колоризация.

#### Введение

Появление первых черно-белых изображений и фильмов было важным шагом человечества к документированию визуальной информации. Однако, как только воплотился в жизнь захват цветных изображений, возникла задача создания цветных версий уже существующих черно-белых материалов. Процесс добавления и распространения цвета в изображениях и видео в оттенках серого называется колоризацией. Прежде всего, колоризация изображений применяется для улучшения визуального качества черно-белых изображений и видео, цветные версии которых не существуют или находятся в плохом состоянии, но также колоризация применима в медицинских изображениях для разметки областей интереса, матировании изображений, дизайне интерьера, средствах моделирования, например, грима и макияжа и в других задачах.

Процесс колоризации фотографий, осуществляемый вручную, возник ненамного позже появления самой фотографии. Существуют примеры ручной колоризации изображений, датируемые 1842 годом, хотя возможны и более ранние примеры. Ручной процесс колоризации фильмов также появился достаточно давно. Так, в начале 1900-х годов это практиковала французская компания Pathe, а также ручной подход колоризации использовался в случае диафильмов в 1930-х годах. Полностью ручным был, например, процесс колоризации кадров космической программы Аполлон в 1960-х, что потребовало колоссального объема работы. Компьютерный процесс для колоризации впервые был предложен в 1970 году для добавления цвета в черно-

Задача создания цветных версий уже существующих чернобелых изображений и видео возникла сразу же после того, как научились снимать и показывать цветные изображения и видео. Процесс добавления и распространения цвета в изображениях и видео в оттенках серого называется колоризацией. И хотя наряду с развитием методов цифровой обработки изображений и видео колоризация также постепенно автоматизируется, для ее осуществления до сих используется много ручного труда. В данной работе рассматривается задача распространения цвета в видео по одному цветному кадру, который может быть получен путем ручной колоризации либо путем применения методов автоматической колоризации изображений. Для решения этой задачи рассмотрено два типа нейросетевых методов: первые основаны на предсказании сепарабельных фильтров, вторые используют оптический поток для распространения цвета между кадрами. Большинство из рассмотренных методов показали высокое качество работы на тестовых выборках с сохранением качества при колоризации на большое количество кадров вперед.

белые фильмы, тогда же и был сформулирован термин колоризация [1].

Наряду с развитием методов цифровой обработки изображений и видео колоризация изображений и видео также постепенно автоматизируется для упрощения ручного труда. Колоризация привлекает большое внимание со стороны исследователей в области компьютерного зрения, и новые технологии, направленные на сокращение необходимого числа человеческих усилий для колоризации изображений и видео, все еще продолжают разрабатываться. В настоящее время колоризация попрежнему считается достаточно сложной проблемой, требующей произвольных высокоуровневых знаний относительно содержания изображений и видео для успешного решения задачи. Поэтому для колоризации до сих используется много ручного труда, а сам процесс осуществ-

ляется посредством использования плагинов к ПО по обработке изображений и видео общего назначения, таких как Adobe Photoshop, Nuke и Adobe After Effects.

В данной работе рассматривается задача распространения цвета с одного цветного кадра видеопоследовательности на все остальные черно-белые кадры. В работе предложены новые нейросетевые методы распространения цвета в видео. В дальнейшем входные данные о цвете можно дополнительно сократить за счет использования методов по распространению цвета внутри кадра и сначала распространять введенные пользователем «мазки», либо не использовать входные данные о цвете вообще и раскрашивать исходный кадр с помощью методов автоматической колоризации изображений.

#### Обзор методов колоризации

В целом постановка задачи колоризации изображений или видео является некорректной, так как эта задача не обладает единственным решением, отвечающим условию высокого визуального качества (например, без уменьшения визуального качества в разные цвета может быть окрашена одежда). Для получения удовлетворительных результатов колоризации было предложено две категории методов: распространение цвета, управляемое пользователем, и автоматическая колоризация.

Методы распространения цвета, управляемого пользователем [2-5], требуют ввода цветных «мазков» поверх изображения в оттенках серого и распространяют полученные цвета по всему изображению, решая глобальную задачу оптимизации. Эти методы могут достичь впечатляющих результатов колоризации, но часто требуют интенсивного взаимодействия с пользователем (иногда число взаимодействий достигает 50 для изображений небольшого размера), так как каждая область изображения с различными цветами должна быть явно отмечена пользователем. Пример такого изображения с входными данными пользователя представлен на рис. 1. Поскольку данные методы полностью полагаются на ввод от пользователей, возникает необходимость указывать цвета даже для областей, в которых цветовая неопределенность не возникает, например, зеленая растительность или синее небо. И хотя методы распространения цвета, управляемого пользователем, могут быть легко адаптированы для колоризации видео путем определения соседних пикселей в оптимизационной функции не только в пространственных координатах, но и во времени за счет использования алгоритмов оптического потока, данная модификация потребует от пользователя колоссальный объем входных данных, так как для каждого кадра придется вручную определить большое число цветных областей. В противном случае, при малом объеме входных данных, результат колоризации будет деградировать со временем, как показано на рис. 2.

Чтобы избавиться от проблемы использования пользовательских данных при решении данной задачи, ранние методы автоматической колоризации раскрашивали изображения путем обучения цветовых признаков по одному или нескольким шаблонным цветным изоб-

ражениям с похожим семантическим содержанием [6-8]. К сожалению, иногда может потребоваться очень много времени для поиска подходящего шаблонного цветного изображения, либо такого изображения может попросту не быть, поэтому данные методы не применялись на практике. Однако за последние несколько лет большую популярность и повсеместное использование в области компьютерного зрения обрели сверточные нейронные сети, существенно улучшив качество работы методов в таких задачах, как классификация изображений, обнаружение объектов, семантическая сегментация и других. Поскольку сверточные нейронные сети обладают способностью обучаться высокоуровневым признакам в изображениях, они также были использованы и для автоматической колоризации [9-11], значительно улучшив показатели данных методов. Нейросетевые методы автоматической колоризации способны генерировать правдоподобные цветные изображения в большинстве случаев (пример удачной и неудачной колоризации с помощью метода [11] представлен на рис. 3).





Рис. 1. Слева представлен пример изображения в оттенках серого с введенными пользователем цветами. Справа — результат колоризации с помощью метода [2]. Как можно видеть, для получения качественной колоризации одного изображения небольшого размера с помощью оптимизационного метода необходимо большое количество взаимодействий с пользователем.



Рис. 2. Пример работы алгоритма [2] для колоризации видео при предоставлении цвета для первого кадра.
Результат работы отчетливо показывает, что использование такого алгоритма

без дополнительных данных для следующих кадров приводит к потере качества колоризации для этих кадров

Однако эти методы обладают рядом существенных недостатков. Ввиду того, что различные объекты могут обладать несколькими подходящими цветами, результаты работы данных методов могут содержать цвета или стили, которые пользователи не ожидали либо хотели бы изменить. Например, пользователям может понадобиться раскрасить гору в зеленый цвет под весенний период времени, но алгоритм представит желтую ко-

лоризацию горы под осенний период времени. Также методы [9-11] решают только задачу колоризации изображений, и для их применения для видео необходимо существенно изменить нейросетевую архитектуру, а также процесс обучения. При использовании сверточной нейронной сети, обученной для колоризации изображений, на видео возможно возникновение сильной нестабильности результатов во времени.





Рис. 3. Пример удачной (слева)

и неудачной (справа) колоризации с помощью метода [11]. В большинстве случаев неудачная колоризация связана с отсутствием хороших аналогов в обучающей выборке

Но недавно начали появляться нейросетевые методы для распространения данных в видео, в том числе цветовых [12, 13]. Данные методы позволяют распространить цвет с одного предоставленного цветного кадра, который может быть получен путем колоризации изображения с помощью описанных выше алгоритмов либо ручной раскраски.

### Предложенные нейросетевые методы Описание входных данных

Видеопоследовательность рассматривается в цветовом пространстве YUV. В качестве входных данных подаются изображения в оттенках серого, соответствующие каналу Y, а также цветовые каналы (U и V) для одного кадра сцены для дальнейшего распространения цвета (данных каналов U и V) с этого кадра на остальные. Распространение цвета на следующие кадры осуществляется по-разному для разных предложенных нейросетевых методов — по-разному учитывается движение в кадрах: с помощью предсказания сепарабельных фильтров непосредственно в сети либо с использованием оптического потока. Далее приведены описания разработанных нейросетевых архитектур.

# Нейросетевые методы, предсказывающие сепарабельные фильтры для распространения цвета

В данных методах в качестве базовой нейросетевой архитектуры была выбрана сеть, предсказывающая сепарабельные фильтры для распространения цвета с одного кадра на другой. Аналогичный подход впервые был использован в задаче интерполяции кадров видео, в котором для получения промежуточных кадров сетью предсказываются сепарабельные фильтры как для первого, так и для второго кадра, осуществляется свертка

соответствующих кадров с предсказанными свертками и результаты сверток складываются [14]. Данный подход был адаптирован на задачу распространения цвета с одного кадра на другой: сеть предсказывает сепарабельный фильтр для первого кадра, а для получения цветовых каналов для второго кадра осуществляется свертка цветовых каналов первого кадра с предсказанным фильтром. Общая схема метода представлена на рис. 4.

В качестве базовой сети была выбрана сеть типа кодировщик-декодировщик, состоящая из набора сверточных блоков. Детали выбранной архитектуры представлены в табл. 1.

Для улучшения качества работы базовой сети в нее был добавлен учет границ за счет использования дифференцируемых фильтров. Для этого были использованы фильтры domain transform [15] и fast global smoother [16]. Дифференцируемость фильтра domain transform была продемонстрирована в работе [17], в которой он был использован для постобработки результатов в нейронной сети для семантической сегментации изображений. Аналогичным свойством обладает и фильтр fast global smoother. Для доказательства рассмотрим этот метод. Данный фильтр решает следующую оптимизационную задачу для 1D сигнала (то есть для каждой строки и каждого столбца изображения):

$$\arg\min_{u} \sum_{x} \left( (u_x - f_x)^2 + \lambda \sum_{i \in N(x)} (u_x - u_i) \right)$$

где u – отфильтрованный сигнал, f – фильтруемый

сигнал, 
$$N(x)=\{x-1,x+1\},$$
  $\omega_{x,x-1}=e^{-\sqrt{2}\frac{g_x}{\sigma_c}},$   $\omega_{x,x+1}=e^{-\sqrt{2}\frac{g_x}{\sigma_c}},$   $\omega_{x,x+1}=e$ 

Решение данной задачи оптимизации осуществляется в два прохода:

1. Прямой проход:

$$\begin{split} \tilde{c}_x &= \frac{c_x}{b_x - \tilde{c}_{x-1} a_x}, \, \tilde{c}_0 = \frac{c_0}{b_0}, \, \tilde{f}_x = \frac{f_x - \tilde{f}_{x-1} a_x}{b_x \tilde{c}_{x-1} a_x}, \, \tilde{f}_0 = \frac{f_0}{b_0}, \\ a_x &= -\lambda \omega_{x,x+1}, \, b_x = 1 + \lambda (\omega_{x,x-1} + \omega_{x+1,x}), \, c_x = -\lambda \omega_{x+1,x}, \\ a_0 &= 0, \, c_{w-1} = 0, \end{split}$$

где w – размерность 1D сигнала.

2. Обратный проход:

$$u_{x} = f_{x} - \tilde{c}_{x} u_{x+1}, \ u_{w-1} = f_{w-1}.$$

Покажем дифференцируемость обратного прохода. При обратном распространении ошибки с последующих слоев нейронной сети приходит  $\dfrac{\partial L}{\partial u_x}$  . Тогда для обрат-

ного прохода получаем следующие производные:

$$\begin{split} &\frac{\partial L}{\partial \tilde{f}_x} \leftarrow \frac{\partial L}{\partial u_x}, \quad \frac{\partial L}{\partial \tilde{c}_x} \leftarrow \frac{\partial L}{\partial u_x} \cdot u_{x+1}, \quad \frac{\partial L}{\partial u_{x+1}} \leftarrow \frac{\partial L}{\partial u_{x+1}} - \frac{\partial L}{\partial u_x} \cdot \tilde{c}_x. \\ &\text{Аналогично для прямого прохода. C обратного прохода} \\ &\text{приходят } \frac{\partial L}{\partial \tilde{f}_x}, \quad \frac{\partial L}{\partial \tilde{c}_x} \ . \end{aligned}$$

водные для прямого прохода:

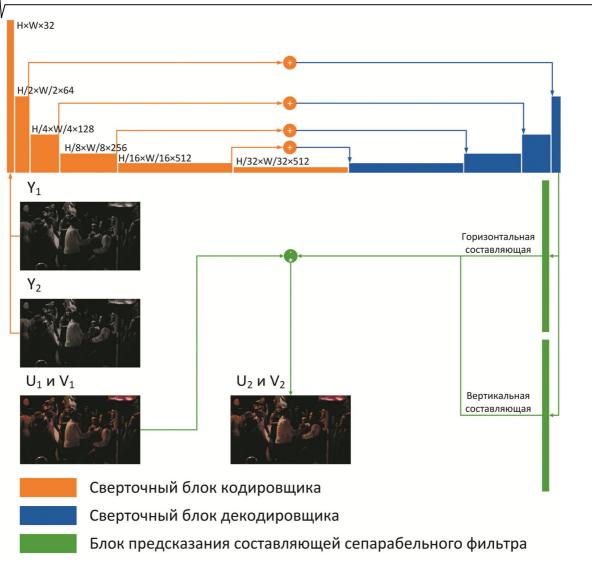


Рис. 4. Схема архитектуры сети, предсказывающей сепарабельные фильтры для распространения цвета с одного кадра на другой

Таблица 1. Конфигурация блоков сверточной нейронной сети для распространения цвета между кадрами видео посредством предсказания сепарабельных фильтров

Слой	Размерность	Шаг	Коэффициент дилатации	Функция активации	
Сверточный блок кодировщика					
Сверточный	3 x 3	1	1	ReLU	
Сверточный	3 x 3	1	1	ReLU	
Сверточный	3 x 3	1	1	ReLU	
Сверточный	3 x 3	1	1	ReLU	
Субдискретизации с выбором среднего	2 x 2	2	1	_	
Сверточный блок декодировщика					
Сверточный	3 x 3	1	1	ReLU	
Сверточный	3 x 3	1	1	ReLU	
Сверточный	3 x 3	1	1	ReLU	
Билинейной интерполяции	_	_	_		
Сверточный	3 x 3		1	ReLU	
Блок предсказания составляющей сепа	рабельного фильтра	ı			
Сверточный	3 x 3	1	1	ReLU	
Сверточный	3 x 3	1	1	ReLU	
Сверточный	3 x 3	1	1	ReLU	
Билинейной интерполяции	_				
Сверточный	3 x 3	1	1	ReLU	
Softmax	<u> </u>	_	_	_	

$$\begin{split} \frac{\partial L}{\partial f_x} &\leftarrow \frac{\partial L}{\partial \tilde{f}_x} \cdot \frac{1}{b_x - \tilde{c}_{x-1} a_x}, \\ \frac{\partial L}{\partial a_x} &\leftarrow \frac{\partial L}{\partial \tilde{f}_x} \cdot \frac{f_x \tilde{c}_{x-1} - \tilde{f}_{x-1} b_x}{(b_x - \tilde{c}_{x-1} a_x)^2} + \frac{\partial L}{\partial \tilde{c}_x} \cdot \frac{c_x \tilde{c}_{x-1}}{(b_x - \tilde{c}_{x-1} a_x)^2}, \\ \frac{\partial L}{\partial c_x} &\leftarrow \frac{\partial L}{\partial \tilde{c}_x} \cdot \frac{1}{b_x - \tilde{c}_{x-1} a_x}, \\ \frac{\partial L}{\partial b_x} &\leftarrow \frac{\partial L}{\partial \tilde{f}_x} \cdot \frac{\tilde{f}_x a_x - f_x}{(b_x - \tilde{c}_{x-1} a_x)^2} + \frac{\partial L}{\partial \tilde{c}_x} \cdot \frac{c_x}{(b_x - \tilde{c}_{x-1} a_x)^2}, \\ \frac{\partial L}{\partial \omega_{x,x-1}} &\leftarrow \left(\frac{\partial L}{\partial b_x} - \frac{\partial L}{\partial a_x}\right) \cdot \lambda, \\ \frac{\partial L}{\partial \omega_{x+1,x}} &\leftarrow \frac{\partial L}{\partial \omega_{x+1,x}} + \left(\frac{\partial L}{\partial b_x} - \frac{\partial L}{\partial c_x}\right) \cdot \lambda, \\ \frac{\partial L}{\partial g_x} &\leftarrow \frac{\partial L}{\partial \omega_{x,x-1}} \cdot \omega_{x,x-1} \cdot \frac{\sqrt{2}}{\sigma_c}, \end{split}$$

$$\begin{split} &\frac{\partial L}{\partial \tilde{f}_{x-1}} \leftarrow \frac{\partial L}{\partial \tilde{f}_{x-1}} - \frac{\partial L}{\partial \tilde{f}_x} \cdot \frac{a_x}{b_x - \tilde{c}_{x-1} a_x}, \\ &\frac{\partial L}{\partial \tilde{c}_{x-1}} \leftarrow \frac{\partial L}{\partial \tilde{c}_{x-1}} + \frac{\partial L}{\partial \tilde{f}_x} \cdot \frac{a_x (f_x - \tilde{f}_x a_x)}{(b_x - \tilde{c}_{x-1} a_x)^2} + \frac{\partial L}{\partial \tilde{c}_x} \cdot \frac{a_x c_x}{(b_x - \tilde{c}_{x-1} a_x)^2}. \end{split}$$

Таким образом, данный фильтр дифференцируем и может быть использован в качестве постобработки для учета границ непосредственно в нейронной сети.

При применении описанных выше фильтров базовый метод модифицируется за счет добавления сверточной сети для предсказания весов фильтров и добавления слоя, реализующего фильтрацию с помощью одного из фильтров. Для предсказания весов фильтров была использована сверточная нейронная сеть для предсказания границ, аналогичная по архитектуре сети [17]. Общая схема модифицированного метода представлена на рис. 5.

Также базовый метод был модифицирован для учета во время обучения нескольких подряд идущих кадров

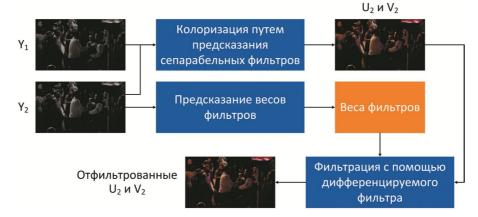


Рис. 5. Схема модифицированного метода, предсказывающего сепарабельные фильтры для распространения цвета с фильтрым дифференцируемых фильтров

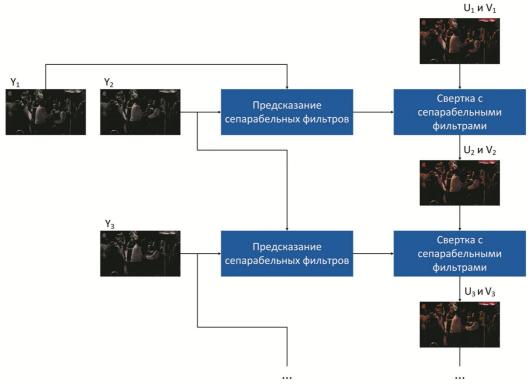


Рис. 6. Схема модифицированного метода, предсказывающего сепарабельные фильтры для распространения цвета с учетом нескольких подряд идущих кадров

в обучающей выборке для повышения итоговой временной стабильности метода и потенциального увеличения качества работы колоризации. Для этого сверточная нейронная сеть, используемая в базовом методе, во время обучения применялась к нескольким подряд идущим кадрам, распространяя цвет с предыдущего кадра на текущий вместо распространения с одного цветного кадра, предоставленного на вход: для каждого текущего кадра для распространения цвета с помощью предсказанных сепарабельных фильтров используется результат колоризации предыдущего кадра. При этом на каждом этапе колоризации очередного кадра используется одна и та же сверточная нейронная сеть. Общая схема модифицированного метода представлена на рис. 6.

# Нейросетевые методы, использующие оптический поток

В данных методах для учета движения между кадрами используется оптический поток. Распространение цвета в видео происходит в два этапа.

- 1. Распространение цвета на соседние кадры осуществляется с помощью оптического потока. В качестве используемого оптического потока выбран метод РWC-Net [18]. Для получения цветов для еще нераскрашенного кадра цвет с входного цветного кадра компенсируется по найденным векторам движения к нераскрашенному кадру с помощью билинейной интерполяции. Но так как найденный оптический поток не везде точен, использование только одного оптического потока для колоризации видео приведет к существенным визуальным артефактам. Для поиска областей, содержащих ошибки оптического потока (например, в областях открытия/закрытия - областях, видимых в одном кадре и невидимых в другом), используется критерий, оценивающий консистентность прямого и обратного оптических потоков путем построения карт разности расстояний между исходными пикселями одного изображения и их отображениями, полученными с помощью перехода сначала по прямому оптическому потоку, а затем по обратному оптическому потоку.
- 2. Распространение цвета внутри кадра и исправление ошибок компенсации осуществляется с помощью сверточной нейронной сети, на вход которой подается изображение в оттенках серого, соответствующее каналу Y, а также цвет (каналы U и V), распространенный на предыдущем шаге с другого кадра, и карта консистентности оптического потока.

Для распространения цвета внутри кадра и исправления ошибок компенсации в первом варианте была использована нейросетевая архитектура GridNet [19], представляющая из себя модификацию архитектуры кодировщик-декодировщик и ранее применявшаяся для задачи семантической сегментации.

Данный метод распространения цвета был также модифицирован для учета нескольких подряд идущих кадров в видео для потенциального улучшения качества работы метода. В отличие от исходного метода модифицированная сеть принимает на вход не только яркость раскрашиваемого изображения (канал Y), цветовые каналы (U и V), полученные распространением цве-

та исходного цветного кадра с помощью оптического потока, и соответствующую карта консистентности оптического потока, но также цветовые каналы (U и V), полученные распространением цвета с предыдущего разукрашенного кадра с помощью оптического потока, и соответствующую этому оптическому потоку карта консистентности. Таким образом, во время обучения рассматриваются не отдельные пары кадров, а несколько подряд идущих кадров. При этом на каждом этапе колоризации очередного кадра используется одна и та же сверточная нейронная сеть. Общая схема модифицированного метода представлена на рис. 7.

Также для задачи распространения цвета были опробованы нейросетевые архитектуры, по-разному учитывающие карту консистентности оптического потока вместо обычного ее подмешивания к входным признакам, как это было сделано в предыдущем методе. Так, для распространения цвета внутри кадра и исправления ошибок компенсации была использована нейросетевая архитектура кодировщик-декодировщик, где вместо обычных сверточных слоев применялись стробированные свертки [20]. Конфигурация данной сети представлена в табл. 2. Аналогично методу, использующему архитектуру GridNet, метод на основе стробированных сверток был модифицирован для учета нескольких подряд идущих кадров в видео. Схема модифицированного метода совпадает со схемой метода модифицированного GridNet (рис. 7) и отличается только используемой для распространения цвета сетью.

#### Обучение разработанных нейросетевых методов

Для обучения предложенных методов была использована обучающая часть одной из версий набора данных Vimeo-90К [21], в которой каждый элемент состоит из семи кадров видеопоследовательности, что позволяет обучать модели, учитывающие несколько подряд идущих кадров. Всего набор данных включает в себя 91701 пример из семи кадров с разрешением 448×256.

В качестве базовой функции стоимости, оптимизируемой во время обучения всех моделей, было использовано среднеквадратическое отклонение предсказанных цветовых каналов  $\hat{C}$  от истинных значений цветовых каналов  $\hat{C}$ :

$$L(\hat{C}, C)_{base} = \sum_{k=1}^{D} (\hat{C}_k - C_k)^2,$$

где D — размерность цветовых каналов. При этом в методах, использующих дифференцируемые фильтры для учета границ, данный функционал вычисляется как для результата работы базовой сети, так и для финального результата, полученного после фильтрации. Пусть  $\hat{C}'$  — промежуточный результат колоризации с помощью базовой сети, предсказывающей сепарабельные фильтры. Тогда для данных методов оптимизируется функционал:

$$L(\hat{C}, \hat{C}', C)_{filtered} = L(\hat{C}', C)_{base} + L(\hat{C}, C)_{base}$$

Также используемая функция стоимости модифицируется для методов, учитывающих несколько подряд идущих кадров, — вычисляется среднеквадратическое отклонение предсказанных цветовых каналов  $\hat{C}^{(i)}$  от

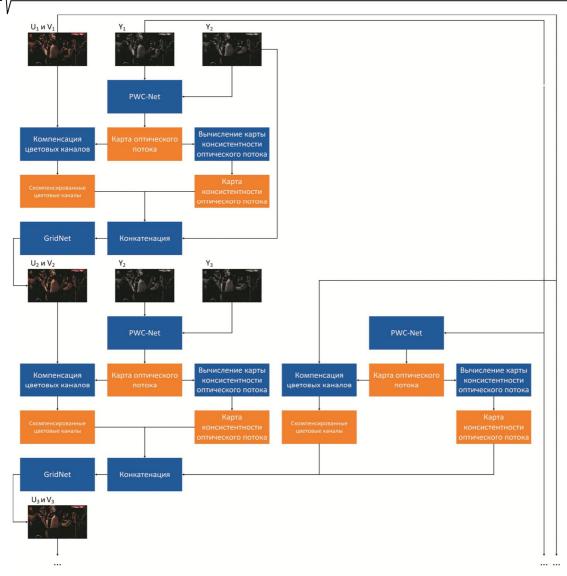


Рис. 7. Схема модифицированного метода для распространения цвета в видео на основе GridNet

Таблица 2. Конфигурация нейронной сети с использованием стробированных сверток

для распространения цвета внутри кадра по результатам компенсации по оптическому потоку

Слой	Размерность	Количество вы-	Шаг	Коэффициент	Функция		
		ходных каналов		дилатации	активации		
Кодировщик	Кодировщик						
Стробированная свертка	3 x 3	32	1	1	Leaky ReLU		
Стробированная свертка	3 x 3	64	2	1	Leaky ReLU		
Стробированная свертка	3 x 3	64	1	1	Leaky ReLU		
Стробированная свертка	3 x 3	128	2	1	Leaky ReLU		
Стробированная свертка	3 x 3	128	1	1	Leaky ReLU		
Стробированная свертка	3 x 3	128	1	1	Leaky ReLU		
Стробированная свертка	3 x 3	128	1	2	Leaky ReLU		
Стробированная свертка	3 x 3	128	1	4	Leaky ReLU		
Стробированная свертка	3 x 3	128	1	8	Leaky ReLU		
Стробированная свертка	3 x 3	128	1	16	Leaky ReLU		
Декодировщик							
Стробированная свертка	3 x 3	128	1	1	Leaky ReLU		
Стробированная свертка	3 x 3	64	1	1	Leaky ReLU		
Билинейной интерполяции	_	_	_	_	_		
Стробированная свертка	3 x 3	64	1	_	_		
Стробированная свертка	3 x 3	32	1	1	Leaky ReLU		
Билинейной интерполяции	_	_	_	_	_		
Стробированная свертка	3 x 3	16	1	1	Leaky ReLU		
Стробированная свертка	3 x 3	2	1	1	Линейная		

истинных значений цветовых каналов  $C^{(i)}$  для каждого рассматриваемого кадра i:

$$L(\hat{C}, C)_{sed} = \sum_{i=1}^{n} \sum_{k=1}^{D} (\hat{C}_{k}^{(i)} - C_{k}^{(i)})^{2},$$

где n — количество учитывающихся кадров в каждом примере, которое было выбрано равным 4.

Перед обучением веса нейронных сетей инициализировались с помощью метода Xavier [22]. Для оптимизации функции стоимости был использован метод Adam [23]. В зависимости от количества параметров в сети и сложности модели использовались следующие параметры обучения: число примеров на каждой итерации обучения — 3-8, количество эпох обучения модели — 20-60, коэффициент скорости обучения — 10<sup>-4</sup>, с уменьшением коэффициента в 10 раз каждую треть от общего количества шагов обучения.

#### Экспериментальная оценка

Тестирование разработанных методов проводилось на тестовой части набора данных Vimeo-90K [21], в которой каждый элемент состоит из семи кадров видеопоследовательности, а также на последовательностях из набора данных АСТ [24], состоящего из 11234 видео, принадлежащих 43 классам действий, с разрешением 448×256. На данных видеопоследовательностях измерялось качество распространения цвета на большое количество кадров, а не только на ближайшие кадры. Также для сравнения были реализованы и обучены аналоги по распространению данных видео — методы TPN [12] и STPN [13]. Во время тестирования вычислялись следующие объективные метрики качества: пико-

вое отношение сигнала к шуму – PSNR, индекс структурного сходства – SSIM.

Результаты тестирования на тестовой части набора данных Vimeo-90К представлены в табл. 3. Как видно из таблицы, лучший результат как по показателю PSNR, так и по показателю SSIM был получен с помощью метода на основе оптического потока и модели GridNet с учетом последовательности кадров. При этом добавление учета последовательности кадров в модель, предсказывающую сепарабельные фильтры, значительно ухудшает качество работы модели и поэтому в дальнейшем не используется.

Результаты тестирования на видеопоследовательностях из набора данных АСТ представлены в табл. 4. В данном случае измерялся средний показатель PSNR для 10, 20, 30 и 40 кадров последовательностей для того, чтобы посмотреть качество колоризации в зависимости от расстояния между раскрашиваемыми кадрами. При тестировании использовались модели, обученные на обучающей части набора данных Vimeo-90K, и дообучения моделей под набор данных АСТ не проводилось. Снова лучшие результаты показала модель на основе оптического потока и модели GridNet с учетом последовательности кадров.

Таким образом, наилучшим качеством обладает модель на основе оптического потока и модели GridNet с учетом последовательности кадров как по сравнению с недавними аналогами, так и с другими разработанными в данной работе методами. Примеры более качественной работы предложенного метода по сравнению с аналогами представлены на рис. 8.

Таблица 3. Результаты тестирования методов распространения цвета в видео на тестовой части набора данных Vimeo-90K

Метод	PSNR	SSIM
TPN [12]	39,2919	0,9760
STPN [13]	39,3702	0,9760
Предсказание сепарабельных фильтров	41,4124	0,9812
Предсказание сепарабельных фильтров + domain transform	41,9002	0,9824
Предсказание сепарабельных фильтров + fast global smoother	41,6647	0,9811
Предсказание сепарабельных фильтров + учет последовательности кадров	38,7839	0,9715
Оптический поток + GridNet	39,0103	0,9737
Оптический поток + GridNet + учет последовательности кадров	41,9163	0,9847
Оптический поток + стробированные свертки	38,9978	0,9749
Оптический поток + стробированные свертки + учет последовательноси кадров	39,9714	0,9799

Таблица 4. Результаты тестирования методов распространения цвета в видео на наборе данных АСТ

Метод	PSNR-10	PSNR-20	PSNR-30	PSNR-40
TPN [12]	35,2908	33,4738	32,4591	31,9222
STPN [13]	35,3388	33,4690	32,4487	31,9014
Предсказание сепарабельных фильтров	35,4241	33,0028	31,7690	31,0834
Предсказание сепарабельных фильтров + domain transform	35,8828	33,4696	32,2324	31,5522
Предсказание сепарабельных фильтров + fast global smoother	36,0216	33,4983	32,2237	31,5199
Оптический поток + GridNet	34,3889	32,1770	31,1329	30,6210
Оптический поток + GridNet + учет последовательности кадров	36,2783	34,1846	32,9208	32,0460
Оптический поток + стробированные свертки	34,3291	32,1950	31,1769	30,6642
Оптический поток + стробированные свертки + учет последовательноси кадров	35,1742	32,7899	31,4909	30,6594

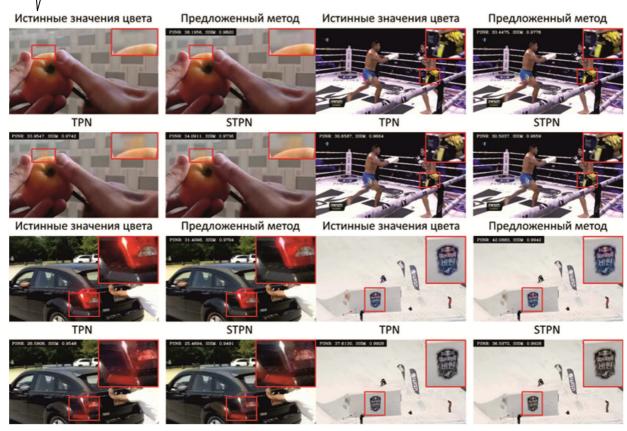


Рис. 8. Примеры высокого качества колоризации с помощью предложенного метода по сравнению с аналогами

#### Заключение

В данной работе предложен ряд методов колоризации видео на основе предсказания сепарабельных фильтров и использования оптического потока для учета движения в кадре. Первый тип методов был также модифицирован путем добавления фильтрации результатов на основе дифференцируемых фильтров, учитывающих границы, дифференцируемость одного из которых была показана в работе. Во втором типе методов были рассмотрены архитектуры GridNet и кодировщикдекодировщик на основе стробированных сверток. Оба типа методов были модифицированы для учета нескольких подряд идущих кадров. Для обучения предложенных методов была использована обучающая часть набора данных Vimeo-90К. Для тестирования была использована тестовая часть набора данных Vimeo-90K, а также сохранение качества результатов со временем было проверено на наборе данных АСТ. Наилучший результат как на тестовой части Vimeo-90K, так и на наборе данных АСТ, по сравнению со всеми предложенными методами и аналогами показал метод на основе оптического потока и модели GridNet с учетом последовательности кадров.

Исследования выполнены при финансовой поддержке РФФИ в рамках научного проекта 19-01-00785 а.

Обучение моделей производилось на вычислительном комплексе IBM Polus факультета вычислительной математики и кибернетики МГУ: http://hpc.cmc.msu.ru/polus.

#### Литература

1. Burns G. Colorization // Museum of Broadcast Communication: Encyclopedia of Television.

- 2. Levin A., Lischinski D., Weiss Y. Colorization using optimization // ACM Transactions on Graphics (ToG). 2004, vol. 23, no. 3, pp. 689-694.
- 3. Horiuchi T., Kotera H. Colorization for monochrome image with texture // Color and Imaging Conference. 2005, no. 1, pp. 245-250.
- 4. Kawulok M., Smolka B. Competitive image colorization // IEEE International Conference on Image Processing (ICIP). 2010, pp. 405-408.
- 5. Kawulok M., Kawulok J., Smolka B. Textural features for scribble-based image colorization // Computer Recognition Systems 4. 2011, pp 269-278.
- 6. Chang Y. et al. Example-based color stylization of images // ACM Transactions on Applied Perception. 2006 vol. 2, no. 3, pp. 322-345.
- 7. Chia A. Y. S. et al. Semantic colorization with internet images // ACM Transactions on Graphics (TOG). ACM, 2011, vol. 30, no. 6, pp. 156.
- 3. Gupta R. K. et al. Image colorization using similar images // Proceedings of the 20th ACM international conference on Multimedia. 2012, pp. 369-378.
- 9. Cheng Z., Yang Q., Sheng B. Deep colorization // Proceedings of the IEEE International Conference on Computer Vision. 2015, pp. 415-423.
- 10. Larsson G., Maire M., Shakhnarovich G. Learning representations for automatic colorization // European Conference on Computer Vision. 2016, pp. 577-593.
- 11. Zhang R., Isola P., Efros A. A. Colorful image colorization // European Conference on Computer Vision. 2016, pp. 649-666.
- 12. Liu S. et al. Learning affinity via spatial propagation networks // Advances in Neural Information Processing Sys-

tems. 2017, pp. 1520-1530.

- 13. Liu S. et al. Switchable temporal propagation network // Proceedings of the European Conference on Computer Vision (ECCV). 2018, pp. 87-102.
- 14. Niklaus S., Mai L., Liu F. Video frame interpolation via adaptive separable convolution // Proceedings of the IEEE International Conference on Computer Vision. 2017, pp. 261-270.
- 15. Gastal E. S. L., Oliveira M. M. Domain transform for edge-aware image and video processing // ACM Transactions on Graphics (ToG). ACM, 2011, vol. 30, no. 4, 69 p.
- 16. Min D. et al. Fast global image smoothing based on weighted least squares // IEEE Transactions on Image Processing. 2014, vol. 23, no. 12, pp. 5638-5653.
- 17. Liu Y. et al. Richer convolutional features for edge detection // Proceedings of the IEEE conference on computer vision and pattern recognition. 2017, pp. 3000-3009.
- 18. Sun D. et al. PWC-Net: CNNs for optical flow using pyramid, warping, and cost volume // Proceedings of the

IEEE Conference on Computer Vision and Pattern Recognition. 2018, pp. 8934-8943.

- 19. Fourure D. et al. Residual conv-deconv grid network for semantic segmentation // arXiv preprint arXiv:1707.07958. 2017.
- 20. Yu J. et al. Free-form image inpainting with gated convolution // Proceedings of the IEEE International Conference on Computer Vision. 2019, pp. 4471-4480.
- 21. Xue T. et al. Video enhancement with task-oriented flow // International Journal of Computer Vision. 2017, pp. 1-20.
- 22. Glorot X., Bengio Y. Understanding the difficulty of training deep feedforward neural networks // Proceedings of the thirteenth international conference on artificial intelligence and statistics. 2010, pp. 249-256.
- 23. Kingma D. P., Ba J. Adam: A method for stochastic optimization // arXiv preprint arXiv:1412.6980. 2014.
- 24. Wang X., Farhadi A., Gupta A. Actions~ transformations // Proceedings of the IEEE conference on Computer Vision and Pattern Recognition. 2016, pp. 2658-2667

### новые книги

### Ремизов С.Л., Рясный Ю.В., Дежина Е.В., Черных Ю.С.

Основы цифровой обработки сигналов / Учебное пособие для вузов — М.: Изд-во «Горячая линия-Телеком», 2019 г. – 492 с.: ил.

Рекомендовано Научно-методическим советом федерального государственного бюджетного образовательного учреждения высшего образования «Сибирский государственный университет телекоммуникаций и информатики» (СибГУТИ) в качестве учебного пособия для студентов, обучающихся по направлениям 11.00.00 — «Электроника, радиотехника и системы связи», 10.00.00 — «Информационная безопасность».

Изложены основы теории дискретных сигналов и цепей, линейной дискретной обработки сигналов, приведены краткие сведения из области аналоговых цепей и сигналов, необходимые для более глубокого понимания изложенного материала. Рассмотрены цифровые избирательные фильтры и методы их проектирования, значительное внимание уделено КИХ-фильтрам с линейной фазой, поскольку они являются основой разработки избирательных фильтров с конеч-

ной импульсной характеристикой. Даны основы теории адаптивной обработки сигналов и рассмотрены вопросы её применения для моделирования и идентификации систем, для коррекции передаточных и частотных характеристик.

Для студентов, обучающихся по направлениям 11.00.00 – «Электроника, радиотехника и системы связи», 10.00.00 – «Информационная безопасность» при изучении курса «Основы цифровой обработки сигналов». Будет полезно читателям, самостоятельно изучающим курс цифровой обработки сигналов.





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 много-каскадных 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 полготовки.



УДК 004.93

# МОДИФИКАЦИЯ VGG-АРХИТЕКТУРЫ В ЗАДАЧАХ УНИМОДАЛЬНОЙ И МУЛЬТИМОДАЛЬНОЙ БИОМЕТРИИ

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

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

Топников А.И., к.т.н., Ярославский государственный университет им. П.Г. Демидова, e-mail: topartgroup@gmail.com;

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

# MODIFICATION OF VGG-ARCHITECTURE FOR UNIMODAL AND MULTIMODAL BIOMETRICS

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

The paper considers a problem of personality recognition using neural network approaches. We developed two algorithms based on the analysis of audio and video data. Both approaches are implemented by using a modification of the VGG convolutional neural network. The first identification method is unimodal. The convolutional neural network CNN-VGGS analyzes the mel-frequency cepstral coefficients of the speech signal. The second algorithm is a multimodal approach based on the original architecture of the bidirectional neural network CNN-VGGMulti. The method classifies a person by merging voice and face analysis. The research results proved the effectiveness of the multimodal algorithm in the problem of personality recognition and this method can be used in real biometric systems.

**Key words:** digital speech processing, digital image processing, speaker identification, face recognition, convolutional neural network, unimodal biometrics, bimodal biometrics.

**Ключевые слова:** цифровая обработка речевых сигналов, цифровая обработка изображений, идентификация диктора, распознавание лиц, сверточная нейронная сеть, унимодальная биометрия, бимодальная биометрия.

#### Введение

В настоящее время существует большое количество сервисов, приложений и услуг, использующих методы биометрической идентификации. Связано это с тем, что данная технология позволяет достаточно точно аутентифицировать пользователя и защитить его персональные данные от несанкционированного доступа. Как правило, это методы анализа по отпечатку пальцев, изображению лица или голосу.

Однако любая унимодальная система имеет свойственный ей ряд ограничений. Биометрия на основе анализа отпечатков пальцев является контактным методом, что уменьшает область практической применимости данной технологии. Системы распознавания пользователя по изображению лица имеют сильную зависимость от уровня освещенности, ракурса, качества фоторегистратора, они также чувствительны к возрастным изменениям и мимике. Система идентификации диктора зависит от эффектов канала передачи информации и микрофона, физиологических особенностей говорящего, акустических свойств окружающей среды [1-7].

В работе описывается разработка метода мультимодальной идентификации личности с использованием лицевой и голосовой биометрии. Такой подход позволяет

Рассматривается задача распознавания личности с применением нейросетевых подходов. Предложены два алгоритма на основе анализа аудио- и видеоданных. Оба подхода реализованы с использованием модификации сверточной нейронной сети VGG-архитектуры. Первый метод идентификации личности является унимодальным. Сверточная нейронная сеть CNN-VGGS анализирует мелчастотные кепстральные коэффициенты речевого сигнала. Второй алгоритм представляет собой мультимодальное решение на базе оригинальной архитектуры двунаправленной нейронной сети CNN-VGGMulti. Метод классифицирует личность, комбинируя результаты анализа голоса и лица. Результаты исследования доказали эффективность мультимодального алгоритма в задаче распознавания личности. Решение может быть использовано для разработки реальных биометрических систем.

создать систему бесконтактного сбора биометрических данных, обладающую высоким уровнем надежности. Использование двух биометрических параметров существенно уменьшает вероятность фальсификации данных [4-6]. Первая часть данной работы посвящена рассмотрению унимодального метода идентификации личности с использованием речевых сигналов. Во второй части рассматривается бимодальный алгоритм, основанный на объединении лицевой и голосовой биометрии.

Нейросетевой подход стал одним из главных инструментов в решении задач детектирования, распознавания и сегментации объектов. В частности, методы и алгоритмы на основе нейронных сетей показывают высокие результаты идентификации людей с использованием цифровых изображений и речевых сигналов [8-13].



Рис. 1. Примеры изображений лиц из набора данных VoxCeleb1

Такие сети также используются в задачах обработки и анализа текстов, в медицине, биохимических исследованиях, робототехнике. Практическую значимость нейросетевых подходов сложно переоценить. Данное исследование также основывается на применении сверточных нейронных сетей.

Целью работы является разработка алгоритмов унимодальной и мультимодальной биометрии на основе анализа речевых сигналов и цифровых изображений лиц.

#### Описание набора данных

Для проведения эксперимента использовалась популярная база VoxCeleb1 — аудиовизуальный набор данных, состоящий из коротких фрагментов человеческой речи и цифровых изображений лиц, извлеченных из видеоинтервью [14]. Процесс подготовки данной базы представляет собой оригинальный метод сбора биометрических данных и состоит из нескольких этапов. Вначале авторы выбрали список известных людей из базы лиц VGG Face, который насчитывает 2622 личности. Вторым этапом для каждого класса выгружались по 50 самых популярных YouTube-видео. Для того, чтобы поиск был более точным, авторы комбинировали имя знаменитости со словом «интервью» при каждом поисковом запросе. Это позволило повысить вероятность того, что конкретный человек действительно присутствует на видео.

Детектирование и выравнивание лиц выполнялись с использованием НОG-детектора и ансамбля регрессионных деревьев [15, 16]. Для трекинга лиц применялись подходы, описанные в [17, 18]. Следующим этапом из видео выделялся аудиосигнал, содержащий речь конкретного человека. Идея заключается в синхронизации движения рта на видео и в речи. Тем самым можно сказать, какому именному лицу принадлежит образец речевой активности. Для этого использовалась нейронная сеть SyncNet [19]. Такой подход способен фильтровать дубляж и закадровый голос. На последнем этапе выполнялась верификация детектированных лиц. В качестве классификатора применялась сеть CNN VGG Face, обученная на наборе данных VGG Face [11].

Тестовая база VoxCeleb1 содержит записи речи спикеров, охватывающих широкий спектр различных национальностей, стиля произношения, профессий и возраста. Набор включает более 150000 речевых сигналов для 1251 класса. Помимо речи база VoxCeleb1 содержит набор цифровых изображений лиц, которые были детектированы и вырезаны в процессе обработки видеороликов (рис. 1). Суммарное количество изображений составляет более 1,2 млн. Изображения обладают следующими свойствами: лица имеют разный угол поворота/наклона головы, цвет лица/волос, наличие/отсутствие очков/бороды и усов; съемка с различными сценами и степенью освещенности [14].

Благодаря наличию хорошо структурированной и размеченной базы данных VoxCeleb1, состоящей из цифровых изображений лиц и речевых сигналов, открывается возможность разработки мультимодальной системы идентификации на основе двух биометрических параметров — лица и голоса. Поскольку набор является достаточно крупным и требовательным к вычислительным ресурсам, количество классов было уменьшено с 1251 до 50.

#### Предобработка данных и описание архитектур сетей

В табл. 1 указан объем исследуемой части аудиовизуальной базы VoxCeleb1. Речевые сигналы представлены в формате wav с частотой дискретизации 16 кГц и уровнем квантования по амплитуде 16 бит. Для того чтобы преобразованные данные можно было компоновать в массивы одной размерности, все речевые сигналы должны быть одной длительности. Из каждого речевого сигнала выделялся фрагмент длительностью в 3 секунды согласно методике, изложенной в [14]. Фрагменты выбирались по принципу случайного выделения из оригинального сигнала.

Таблица 1. Статистика анализируемой части набора данных VoxCeleb1

	Обучение	Проверка	Тест	Общее
Изображения	43243	2382	2447	48072
Речевые	5730	271	322	6323
сигналы				

Оригинальные звуковые данные представляют собой изменение амплитуды колебаний во времени, что является не самой информативной формой представления речевого сигнала, поэтому в работе использовались мел-частотные кепстральные коэффициенты (МЧКК). При решении задачи автоматического распознавания речи, как правило, применяется 40-80 фильтров. В работе сформирован банк из 80 треугольных мел-фильтров (рис. 2). В итоге каждый речевой сигнал представлялся матрицей размером 80х301 [20, 21].

Для повышения обобщающей способности обучаемых моделей нейронных сетей применялся метод синтетического увеличения данных. Речевые сигналы подвергались искажениям и преобразованиям: добавление аддитивного белого гауссовского шума; смещение и растяжение сигнала по времени; изменение высоты тембра; использование эффекта реверберации, позволяющего искусственным образом изменять свойства аудиосигнала, меняя представления о масштабе и глубине акустической сцены; применение медианной фильтрации для разделения гармонических и перкуссионных компонент сигнала (HPSS, Harmonic Percussive Source Separation).

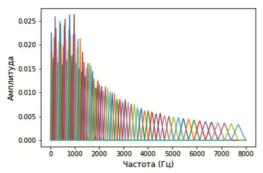


Рис. 2. Банк из 80 треугольных мел-фильтров

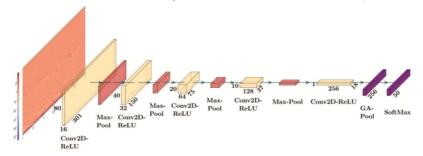
Дополнительно, для повышения вариативности данных, использовался открытый набор звуковых сигналов Urban Sound Dataset (UrbanSound8K). Набор включает 8732 записи, каждая длиной менее 4-х секунд, представляющие собой характерные для города звуки: работа кондиционера, автомобильный гудок, игра детей, лаянье собаки, работа двигателя на холостом ходу, выстрел пистолета, работа отбойного молотка, сирена,

звуки уличной музыки. Речевые сигналы случайным образом смешивались с сигналами из набора UrbanSound8K [22-25].

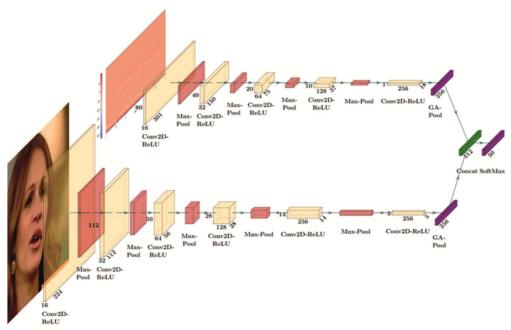
В настоящее время публикуется большое количество интересных исследований на основе анализа наборов данных VoxCeleb1 и VoxCeleb2. Высокие результаты при проведении экспериментов получаются с использованием крупных архитектур: ResNet18, ResNet34, ResNet50 [14, 26]. Для решаемой задачи они не подойдут, поскольку исследуется лишь небольшая выборка набора данных VoxCeleb1, которая составляет менее 5 % от общего объема. Если применять очень глубокие архитектуры к такому малому набору данных, очевидно, что сеть будет переобучаться и обладать слабой обобщающей способностью. Поэтому для данного исследования спроектированы более компактные сверточные нейронные сети.

На рис. 3 представлена CNN-VGGS — архитектура сети, используемая для создания унимодальной системы распознавания личности на основе анализа речевого сигнала. Сверточная нейронная сеть является очень компактной, поскольку содержит менее 0,3 млн. весовых коэффициентов. Для сравнения: сеть ResNet18 содержит более 11 млн. весов, а сеть ResNet50 — более 25 млн. весов.

Для решения задачи мультимодальной идентификации личности по двум биометрическим параметрам спроектирована архитектура CNN-VGGMulti (рис. 4).



Puc. 3. Архитектура сверточной нейронной сети CNN-VGGS



Puc. 4. Архитектура сверточной нейронной сети CNN-VGGMulti

Данная сеть имеет два входа: один для приема цифровых изображений лиц размером 224х224х3, другой — для МЧКК-представления речевых сигналов. Каждый из потоков сети CNN-VGGMulti имеет слой глобального усреднения (GA-Pool, Global Average Pooling). На выходе этих слоев формируются векторы одинаковой размерности по 256 значений, которые далее объединяются с использованием слоя конкатенации Concat в один общий вектор размерностью 512. Это делается для того, чтобы признаки, формируемые разными модальностями, имели одинаковое, равновесное влияние на итоговый результат классификации. Сеть CNN-VGGMulti также является относительно компактной и содержит 0,95 млн. весовых коэффициентов.

# Результаты исследования унимодального алгоритма на основе голосовой биометрии

Опишем процесс обучения нейронной сети CNN-VGGS, а также проведем анализ результатов тестирования. В качестве метода оптимизации весовых параметров применялся Adam (Adaptive Moment Estimation) [27]. В процессе обучения устанавливались следующие гиперпараметры: скорость сходимости алгоритма оптимизации 0,001, размер батча 32, количество эпох 100. На рис. 5 представлен процесс обучения сети с использованием метрики оценки качества ассигасу. Из результатов видно, что на тренировочном наборе данных доля правильных ответов составляет 99,84 %, однако на проверочном множестве оценка имеет значение в 60,89 %, а на тестовом множестве 51,55 %. Такой результат свидетельствует о переобучении модели и ее низкой обобщающей способности. Также видно, что кривая обучения выходит на плато и увеличение количества эпох не дает существенного улучшения в работе.

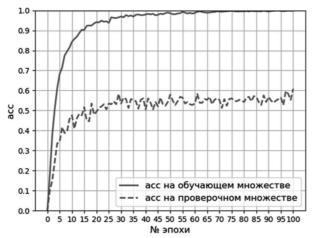
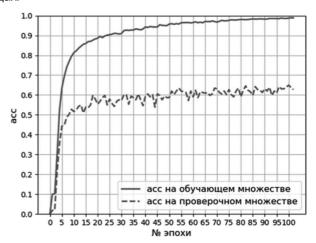


Рис. 5. Кривые изменения доли правильных ответов в процессе обучения сети CNN-VGGS

Одним из распространенных методов борьбы с переобучением является использование одной из разновидностей регуляризации — прореживание слоев. Проведен ряд экспериментов с применением данного подхода, однако получить каких-либо улучшений в точности работы сети не удалось.

Еще одним общепринятым методом повышения обобщающей способности модели является увеличение обучающего набора данных. Новые образцы можно ис-

кать в открытых источниках и сети Интернет или синтетически сгенерировать с использованием различных типов преобразований речевого сигнала. В данном исследовании звуковые сигналы подвергались искажениям и изменениям, описанным ранее. На рис. 6 представлен процесс обучения сети на сгенерированных данных. Применение метода искусственного аугментирования позволило немного улучшить оценку на проверочном и тестовом множествах — 64,95 % и 57,76 % соответственно. Однако модель по-прежнему обладает слабой обобщающей способностью и не способна с высокой точностью определить диктора на новых образнах.



Puc. 6. Кривые изменения доли правильных ответов в процессе обучения сети CNN-VGGS на аугментированных данных

Поскольку метод синтетического увеличения аудиоданных не дал качественного улучшения точности при распознавании личности было решено добавить еще один биометрических параметр, который бы явным образом характеризовал человека. В результате был реализован метод идентификации на основе комбинированного анализа речевых сигналов и цифровых изображений лиц.

### Результаты исследования мультимодального алгоритма на основе лицевой и голосовой биометрии

Поскольку унимодальный метод распознавания диктора не продемонстрировал высоких результатов на проверочном и тестовом множествах был разработан мультимодальный алгоритм. Для этого реализована архитектура CNN-VGGMulti, состоящая из двух веток, имеющих один общий выход. Ветка для анализа речевых данных представляет собой сверточную нейронную сеть CNN-VGGS. Ветка для анализа изображений лиц также представляет собой небольшую сверточную нейронную сеть. Каждая из веток формирует 256-мерный вектор на основе соответствующего ей входа с использованием МЧКК-представления речевого сигнала или цифрового изображения лица. Далее эти вектора объединяются в общий вектор размерностью 512. На входе сети CNN-VGGMulti имеются тензоры размером 80x301x1 и 224x22x3. В качестве метода оптимизации весовых параметров, также, как и для унимодального

алгоритма, применялся метод оптимизации весовых параметров Adam. В процессе обучения устанавливались следующие гиперпараметры: скорость сходимости алгоритма оптимизации 0,001, размер батча 8, количество эпох 100. На рис. 7 представлен процесс обучения мудьтимодального алгоритма на основе сверточной нейронной сети CNN-VGGMulti.

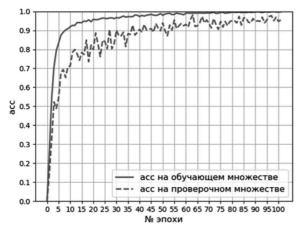


Рис. 7. Кривые изменения доли правильных ответов в процессе обучения сети CNN-VGGMulti

Мультимодальная модель продемонстрировала высокие результаты работы. Оценка на обучающей выборке составила 99,88%, на проверочном и тестовом наборах данных — 98,11% и 97,19% соответственно. Использование комбинированного подхода на основе анализа речевой и лицевой биометрии позволило значительно повысить обобщающую способность классификатора. Обученная модель делает предсказание, основываясь одновременно на двух независимых особенностях человека. Результаты исследования могут быть использованы для проектирования коммерческих систем распознавания личности.

#### Заключение

В работе рассмотрен вопрос классификации личности с использованием методов унимодальной и бимодальной биометрии. В качестве источника аудиовизуальной информации выбран современный набор данных VoxCeleb1. Для проведения исследования из данной базы выделены 50 уникальных классов. Вся база VoxCeleb1 в исследовании не применялась, что позволило более динамично проводить опытные работы на всех этапах эксперимента. Анализ речевых сигналов основывался на выделении мел-частотных кепстральных коэффициентов.

В ходе работы спроектированы две архитектуры сверточных нейронных сетей. Для унимодальной идентификации личности на основе только речевых сигналов применялась сверточная нейронная сеть CNN-VGGS. Исследование показало, что обученная модель имеет слабую обобщающую способность, показав на тестовом множестве точность 51,55 %. Для борьбы с переобучением применялся метод прореживания слоев и искусственная аугментация данных, однако существенным образом это не повлияло на качество работы. Модель, обученная на новых синтезированных данных, показала на тестовом множестве точность 57.76 %.

Для повышения точности идентификации реализован мультимодальный алгоритм распознавания на основе анализа речевых сигналов и цифровых изображений лиц. Разработана архитектура двунаправленной сверточной нейронной сети CNN-VGGMulti. Мультимодальное решение продемонстрировало высокий уровень точности на проверочном и тестовом наборах данных — 98,11 % и 97,19 % соответственно. Результаты работы бимодального алгоритма свидетельствуют о возможности практической применимости данного подхода в реальных коммерческих системах распознавания личности.

На следующем этапе исследования планируется проведение эксперимента с использованием всего набора данных VoxCeleb1. Для решения задачи распознавания личности будет применен метод комбинирования биометрических параметров, описанный в данной статье, но с использованием уже более глубоких топологий сетей ResNet18 и ResNet34. Дополнительно будут представлены новые методы объединения биометрических параметров.

Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 19-37-90158.

#### Литература

- 1. Хайкин С. Нейронные сети: Полный курс. Пер. с англ. Н.Н. Куссуль, А.Ю. Шелестова. 2-е изд. М.: Издательский дом Вильямс, 2008. 1103 с.
- 2. 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.
- 3. 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.
- 4. 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 // Цифровая обработка сигналов. 2017. № 3. С. 45-49.
- 5. Khryashchev V., Topnikov A., Stefanidi A., Priorov A. Bimodal person identification using voice data and face images. In Proceedings SPIE 11041, Eleventh International Conference on Machine Vision, 2018, Web: https://doi.org/10.1117/12.2523138.
- 6. Stefanidi A., Topnikov A., Tupitsin G., Priorov A. Application of convolutional neural networks for multimodal identification task, 26th Conference of Open Innovations Association FRUCT, 423-428, (2020); doi: 10.23919/FRUCT 48808.2020.9087458
- 7. Tupitsin G., Topnikov A., Priorov A. Two-step noise reduction based on soft mask for robust speaker identification, In Proceedings 18th Conference of Open Innovations Association FRUCT. 2016. pp. 351-356.
- 8. Ault S.V., Perez R.J., Kimble C.A., Wang J. On Speech Recognition Algorithms. International Journal of Machine Learning and Computing, vol. 8, no. 6, 2018, pp. 518-523.

- 9. Bunrit S., Inkian T., Kerdprasop N. Text-Independent Speaker Identification Using Deep Learning Model of Convolution Neural Network. International Journal of Machine Learning and Computing, Vol. 9, no. 2, 2019, pp. 143-148.
- 10. 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.
- 11. Simonyan K., Zisserman A. Very deep convolutional networks for large-scale image recognition. In International Conference on Learning Representations, 2015, Web: https://arxiv.org/abs/ 1409.1556v6.
- 12. Parkhi O.M., Vedaldi A., Zisserman A. Deep face recognition. In Proceedings British Machine Vision Conference, 1, 41.1-41.12 10.5244, 2015, pp. 29-41.
- 13. Sun Y., Ding L., Wang X., Tang X. DeepID3: Face recognition with very deep neural networks, 2015, Web: https://arxiv.org/ abs/1502.00873.
- 14. Nagrani A., Chung J.S., Zisserman A. VoxCeleb: a large-scale speaker identification dataset, 2017, Web: https://arxiv.org/abs/ 1706.08612v2.
- 15. King D.E. Dlib-ml: A machine learning toolkit. The Journal of Machine Learning Research, vol. 10, pp. 1755-1758, 2009.
- 16. Kazemi V., Sullivan J. One millisecond face alignment with an ensemble of regression trees. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1867-1874, 2014.
- 17. Chung J. S., Zisserman A. Lip reading in the wild. In Proceedings of the Asian Conference on Computer Vision, 2016.

- 18. Everingham M., Sivic J., Zisserman A. Taking the bite out of automatic naming of characters in TV video. Image and Vision Computing, vol. 27, no. 5, 2009.
- 19. Chung J. S., Zisserman A. Out of time: automated lip sync in the wild. In Workshop on multi-view lip-reading, ACCV, 2016.
- 20. Logan B. Mel frequency cepstral coefficients for music modeling. In International Symposium Music Information Retrieval, 2000.
- 21. Kopparapu S.K., Laxminarayana M. Choice of Mel filter bank in computing MFCC of a resampled speech. 2010. pp. 121-124, doi: 10.1109/ISSPA.2010.5605491.
- 22. Salamon J., Bello J.P. Deep Convolutional Neural Networks and Data Augmentation for Environmental Sound Classification. In IEEE Signal Processing Letters, vol. 24, no. 3, pp. 279-283, March 2017, doi: 10.1109/LSP.2017. 2657381.
- 23. Park D.S., Chan W., Zhang Y., Chiu C.C., Zoph B., Cubuk E.D., Le Q.V. SpecAugment: A Simple Data Augmentation Method for Automatic Speech Recognition. 2019.
- 24. Ko T., Peddinti V., Povey D., Khudanpur S. Audio augmentation for speech recognition. In INTERSPEECH-2015, 2015. pp. 3586-3589.
- 25. Salamon J., Jacoby C., Bello J.P. A Dataset and Taxonomy for Urban Sound Research. 22nd ACM International Conference on Multimedia, Orlando USA, Nov. 2014.
- 26. Chung J.S., Nagrani A., Zisserman A. VoxCeleb2: Deep Speaker Recognition. In Proceedings Interspeech, 2018, pp. 1086-1090.
- 27. Kingma D.P., Ba J. Adam: A Method for Stochastic Optimization. 2017, Web: https://arxiv.org/ abs/1412. 6980v9.

#### новые книги

# Подкорытов А.Н.

Методы местоопределения потребителя в глобальных навигационных спутниковых системах / Учебное пособие для вузов — М.: Изд-во «Горячая линия-Телеком», 2020 г. — 136 стр.: ил.

Приведена классификация основных методов местоопределения в глобальных навигационных спутниковых системах. Рассмотрены абсолютные и относительные методы местоопределения различной точности и оперативности по измерениям систем ГЛОНАСС и GPS, в том числе с использованием систем широкозонной дифференциальной коррекции. Описано местоопределение как статического, так и динамического потребителя. Значительное внимание уделено высокоточному абсолютному местоопределению. Изложены теоретические и методические материалы к лабораторным работам по местоопределению потребителя в глобальных навигационных спутниковых системах.

Для студентов вузов, обучающихся по укрупненным группам специальностей и направлениям подготовки: 24.00.00 – «Авиационная и ракетно-космическая техника», 11.00.00 – «Электроника, радиотехника и системы связи». Будет полезно аспирантам и специалистам соответствующих направлений и научной специальности ВАК 05.12.14.

# Уважаемые коллеги!

Для тех, кто не успел оформить подписку на второе полугодие 2020 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.dspa.ru.

Справки по телефону: (+7 903) 201-53-33 (Самсонов Геннадий Андреевич). E-mail: rntores@mail.ru



УДК 621.397

# ПОВЫШЕНИЕ КАЧЕСТВА СОПРОВОЖДЕНИЯ ОБЪЕКТОВ НА ОСНОВЕ АНАЛИЗА МИНИМУМОВ КРИТЕРИАЛЬНОЙ ФУНКЦИИ

Трофимов Д.В., начальник лаборатории АО «РИФ», e-mail: samael1978@rambler.ru.

# IMPROVING THE QUALITY OF OBJECT TRACKING BASED ON THE ANALYSIS OF MINIMA OF THE CRITERION FUNCTION

### Trofimov D.V.

The problem of tracking failure for correlation automatic video tracking algorithms is actual. The conditions under which such failures occur are considered. To reduce the number of tracking failures under the considered conditions, a procedure for analyzing local minima of the criterion function has been developed, based on three features: the value of the minimum, the minimum sharpness and the distance from minimum to the center of the definition area. An example is shown that confirms the effectiveness of the proposed criteria.

Key words: computer vision, object tracking, correlation algorithms, features, analysis of minima.

**Ключевые слова:** техническое зрение, сопровождение объектов, корреляционные алгоритмы, признаки, анализ минимумов.

#### Введение

Современные системы технического зрения, используемые в различных областях техники, решают множество задач. Одной из таких задач является сопровождение объектов на видеопоследовательностях [1-4].

Для решения данной задачи известны различные подходы, каждый из которых представлен широким набором алгоритмов [5-7]. При этом для сопровождения применяются также методы обнаружения объектов с последующим измерением их координат. В работе [5] выделены следующие классы методов сопровождения объектов на видеопоследовательностях:

- корреляционные методы;
- методы статистической сегментации;
- методы пространственной фильтрации;
- методы пространственно-временной фильтрации.

Несмотря на появление большого количества новых работ по данной тематике, широко распространены методы, основанные на корреляционном сопоставлении текущего кадра с заранее захваченным эталоном (шаблоном), обновляемым в ходе сопровождения. Такие методы представлены широким набором алгоритмов, общий принцип данных алгоритмов можно описать следующим образом.

По стробу изображения  ${f I}$  размером  $S_{\rm IX} \times S_{\rm IY}$  элементов перемещается эталон объекта  ${f R}$  размером  $S_{\rm RX} \times S_{\rm RY}$  элементов. Сдвиги эталона по двум координатам составляют  $x=\overline{1,S_{\rm X}}$  и  $y=\overline{1,S_{\rm Y}}$ , причём максимальные значения сдвига равны  $S_{\rm X}=S_{\rm RX}-S_{\rm IX}+1$  и  $S_{\rm Y}=S_{\rm RY}-S_{\rm IY}+1$  соответственно. Каждому значению сдвига (x,y) ставится в соответствие мера отличия эталона от соответствующего участка изображения F(x,y), называемая в дальнейшем критериальной

Для корреляционных алгоритмов автоматического сопровождения объектов на видеопоследовательностях актуальна проблема срыва сопровождения. Рассмотрены условия, при которых происходят такие срывы. Для уменьшения количества срывов сопровождения в указанных условиях разработана процедура анализа локальных минимумов критериальной функции, основанная на трех признаках: значение минимума, острота вершины и расстояние от центра области определения. Показан пример, подтверждающий эффективность предложенных критериев.

функцией. По координатам минимума критериальной функции определяется положение сопровождаемого объекта относительно строба  ${\bf I}$  .

В данной работе рассмотрен наиболее простой с точки зрения вычислительных затрат способ определения значений критериальной функции, основанный на разностной корреляции [8]:

$$F(x,y) = \sum_{i=1}^{S_{RX}} \sum_{i=1}^{S_{RY}} \left| I_{i+x-1,j+y-1} - R_{i,j} \right|,$$
 (1)

где R и I – соответственно элементы эталона  ${\bf R}$  и строба  ${\bf I}$  .

Для улучшения качества сопровождения при помощи корреляционных алгоритмов используются специальные модификации, описанные, например, в работе [9].

Как правило, в корреляционных алгоритмах для определения положения сопровождаемого объекта используется глобальный минимум критериальной функции. Однако, как показано далее, в некоторых ситуациях истинному положению сопровождаемого объекта соответствует не глобальный, а один из локальных минимумов функции F(x,y).

При сопровождении подвижного объекта на видеопоследовательности возможны ситуации, когда значение средней яркости фона приближается к средней яркости эталона  ${f R}$  :

$$\sum_{i=1}^{S_{RX}} \sum_{j=1}^{S_{RY}} I_{i+x_n-1,j+y_n-1} \approx \sum_{i=1}^{S_{RX}} \sum_{j=1}^{S_{RY}} R_{i,j} , \qquad (2)$$

где  $x_n, y_n$  – координаты локального минимума с поряд-

ковым номером n, не соответствующего истинному положению объекта.

В свою очередь, изменение условий освещения объекта может привести к значительному отличию средней яркости эталона  ${\bf R}$  от средней яркости сопровождаемого объекта, присутствующего в стробе  ${\bf I}$ :

$$\left| \sum_{i=1}^{S_{RX}} \sum_{i=1}^{S_{RY}} I_{i+x_{gt}-1,j+y_{gt}-1} - \sum_{i=1}^{S_{RX}} \sum_{i=1}^{S_{RY}} R_{i,j} \right| >> 0,$$
 (3)

где  $x_{\rm gt}$ ,  $y_{\rm gt}$  – координаты «истинного» минимума, то есть минимума, соответствующего истинному расположению сопровождаемого объекта.

Если эталон и фон не имеют явно выраженной контрастной текстуры, то в описанных ситуациях один или несколько минимумов критериальной функции с координатами  $(x_n, y_n)$  имеют меньшее значение, чем «истинный» минимум с координатами  $(x_n, y_n)$ :

$$F(x_n, y_n) < F(x_{gt}, y_{gt}), \tag{4}$$

Очевидно, что использование глобального минимума для определения координат сопровождаемого объекта в данном случае приведёт к срыву сопровождения. Тем не менее, в точке  $(x_{\rm gt}, y_{\rm gt})$ , как правило, имеется локальный минимум критериальной функции, обусловленный подобием текстур эталона и объекта.

Для того, чтобы такой «истинный» минимум имел меньшее значение, чем иные минимумы, в работе [0] используется штрафная функция, которая учитывает отличие измеренных координат сопровождаемого объекта от прогнозируемых.

В настоящей работе предложена процедура анализа минимумов критериальной функции, основанная на расчёте значений трёх признаков. Показано, что определение «истинного» минимума по данным признакам увеличивает длительность сопровождения по сравнению с аналогичным корреляционным алгоритмом, основанном на выборе глобального минимума.

# Процедура анализа локальных минимумов критериальной функции

Для выбора локального минимума предлагается использовать три признака: значение минимума, острота вершины и расстояние от центра области определения критериальной функции. Полная процедура анализа локальных минимумов критериальной функции включает выделение локальных минимумов, расчёт значений признаков и совместное использование рассчитанных признаков для выбора «истинного» минимума. Данная процедура состоит из шести этапов.

1. Выделяются локальные минимумы критериальной функции в восьмисвязной области. При этом точка с координатами (x,y) считается локальным минимумом, если она является минимальной в окружающей её зоне  $\mathbf{O}$  размером  $3 \times 3$  элемента:

$$F(x, y) = \min(\mathbf{O}). \tag{5}$$

Смещения (x,y) области  ${\bf O}$  изменяются в пределах  $x=\overline{2,S_{\rm X}-1}$  и  $y=\overline{2,S_{\rm Y}-1}$ . Такое исключение краевых пикселей строба  ${\bf I}$  из анализа необходимо не только

для упрощения реализации предлагаемой процедуры, но и для обеспечения устойчивости к минимумам, не являющихся минимальными значениями в восьмисвязной области.

2. Для приведения области значений критериальной функции к интервалу [0;1] Производится нормировка и смещение значений F(x,y):

$$\dot{\mathbf{F}} = \frac{\mathbf{F} - \min(\mathbf{F})}{\max(\mathbf{F}) - \min(\mathbf{F})},\tag{6}$$

где **F** – смещённая и нормированная критериальная функция.

3. Для каждого определённого ранее минимума с порядковым номером  $n=\overline{1,N}$  рассчитывается признак  $A_n$ , основанный на значении критериальной функции в точке  $(x_n,y_n)$ :

$$A_{n} = 1 - w_{A} \dot{F}(x_{n}, y_{n}), \tag{7}$$

где  $w_{\rm A}$  – вес признака A, используемый при мультипликативном объединении трёх признаков.

4. Для каждого минимума рассчитывается расстояние  $d_n$  от точки  $(x_n,y_n)$  до центра области определения  $\mathbf{F}$ , нормированное к половине диагонали этой области (то есть максимально возможному расстоянию):

$$d_n = \frac{\sqrt{\left(x_n - \frac{S_X + 1}{2}\right)^2 - \left(y_n - \frac{S_Y + 1}{2}\right)^2}}{\sqrt{\left(1 - \frac{S_X + 1}{2}\right)^2 - \left(1 - \frac{S_Y + 1}{2}\right)^2}}.$$
 (8)

Значение второго признака рассчитывается на основе расстояния  $d_{\scriptscriptstyle n}$ :

$$B_n = 1 - w_{\rm B} d_n \,, \tag{9}$$

где  $w_{\rm B}$  – вес признака B .

5. Для каждого минимума  $(x_n, y_n)$  рассчитывается коэффициент остроты вершины  $v_n$ :

$$v_{n} = \left(\frac{1}{(2S_{V} + 1)^{2}} \sum_{i=-S_{V}}^{S_{V}} \sum_{j=-S_{V}}^{S_{V}} \dot{F}(x_{n} + i, y_{n} + j)\right) - \dot{F}(x_{n}, y_{n}),$$
(10)

где  $S_{
m V}$  — половина размера окрестности минимума. В данной работе использовалось значение  $S_{
m V}=5$  .

Для тех минимумов  $(x_n,y_n)$ , для которых значения  $x_n+i$  или  $y_n+j$  при  $i,j=\overline{-S_{\rm V},S_{\rm V}}$  выходят за границы области определения критериальной функции  ${\rm F}(x,y)$ , пределы изменения переменных i и j уменьшаются таким образом, чтобы описанного выхода за границы не происходило.

После того как для каждого локального минимума получены значения  $v_n$ , данные значения смещаются и нормируются в ходе преобразования, аналогичному преобразованию (6). Нормированные и смещённые коэффициенты  $\dot{v}_n$  используются для расчёта значений третьего признака:

$$C_n = w_C \dot{v}_n + (1 - w_C) , \qquad (11)$$

где  $w_{C}$  – вес признака C .

6. Рассчитанные признаки объединяются для принятия решения о выборе минимума:

$$T_n = A_n B_n C_n . ag{12}$$

Минимум с наибольшим значением  $T_n$  принимается за «истинный» минимум.

# **Анализ результатов работы предложенной процедуры**

Типовой пример критериальной функции, при которой происходит срыв сопровождения по описанным выше причинам, показан на рисунке.

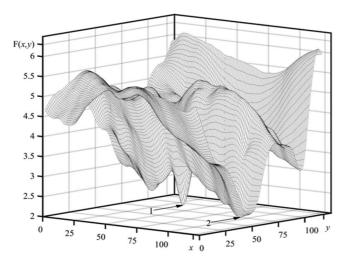


Рис. 1. Пример формы критериальной функции, при которой происходит срыв сопровождения при использовании только глобального минимума Цифрой 1 на данном рисунке отмечен «истинный» минимум, а цифрой 2 — глобальный минимум.

Из анализа рисунка следуют выводы:

Минимум 1 не является глобальным, то есть выбор глобального минимума приведёт к срыву сопровождения.

Данный минимум расположен ближе остальных минимумов к центру области определения критериальной функции, то есть данному минимуму должно соответствовать наибольшее значение признака B.

«Истинный» минимум имеет более острую вершину

по сравнению с другими минимумами, то есть данномуминимуму должно соответствовать наибольшее значение признака B.

Значения критериев  $A,\ B$  и C для различных минимумов, а также их произведение T при равных весах  $w_{\rm A}=w_{\rm B}=w_{\rm C}=1$ . представлены в табл. 1.

В первой строке данной таблицы представлены значения для «истинного» минимума, во второй – для глобального, а далее – для остальных минимумов показанной на рисунке критериальной функции.

Из анализа табл. 1 следует, что «истинный» минимум проигрывает другим минимумам, только по признаку A. При этом помимо глобального минимума (n=2) ещё два минимума (порядковые номера 6 и 7) имеют близкие значения критериальной функции. По двум предложенным признакам B и C «истинный» минимум существенно превосходит остальные минимумы. В результате объединения трёх признаков данное превосходство увеличивается: по значениям произведения T превосходство «истинного» минимума по сравнению с ближайшим по T составляет 0,952 против 0,162, при этом глобальному минимуму соответствует значение  $T_2=0,071$ .

Для оценки качества сопровождения объектов создана программа, позволяющая в полуавтоматическом режиме размечать треки объектов на видеопоследовательности. При помощи данной программы размечены треки семи объектов на семи различных видеопоследовательностях.

Сопровождение данных объектов производилось двумя способами: при помощи предложенной процедуры и при помощи выбора глобального минимума. В каждом кадре рассчитывалась ошибка определения координат объекта, нормированная к половине его диагонали. За срыв сопровождения принималась ситуация, при которой данная ошибка превышала пороговое значение  $\lambda$ . При этом в экспериментальных исследованиях использовались два порога:  $\lambda=0,25$  и  $\lambda=0,5$ . Мерой качества сопровождения в данном эксперименте являлась длительность сопровождения до первого превышения ошибкой порога  $\lambda$ .

Длительности сопровождения различных объектов двумя рассмотренными способами при двух различных порогах  $\lambda$  приведены в табл. 2. (в кадрах).

Таблица 1. Значения описанных признаков для различных минимумов

Порядковый номер	Признак $A_n$	Признак $B_n$	Признак $C_n$	Произведение $T_n$	
минимума п	- "	- "	- "	- "	
1 (истинный)	0,985	0,967	1	0,952	
2 (глобальный)	1	0,380	0,187	0,071	
3	0,207	0,348	0	0	
4	0,277	0,474	0,063	0,008	
5	0,300	0,557	0,066	0,011	
6	0,931	0,533	0,302	0,150	
7	0,926	0,625	0,279	0,162	
8	0,800	0,337	0,203	0,055	
9	0,530	0,503	0,132	0,035	
10	0,544	0,516	0,051	0,014	
11	0,650	0,365	0,197	0,047	
12	0,651	0,378	0,247	0,061	

Таблица 2. Длительности сопровождения различных объектов

Порядковый номер	Длительность сопровождения, кадр.					
трека	Глобальный минимум		Предложенная процедура			
	$\lambda = 0,25$	$\lambda = 0,5$	$\lambda = 0,25$	$\lambda = 0.5$		
1	48	48	225	225		
2	9	9	63	63		
3	25	25	99	100		
4	39	39	40	44		
5	32	32	100	100		
6	87	136	85	139		
7	23	23	45	46		

Из анализа данной таблицы следует, что применение предлагаемой процедуры анализа минимумов критериальной функции приводит к увеличению длительности сопровождения до срыва в среднем в 2,3 раза.

#### Заключение

В ходе данной работы показано, что при сопровождении объектов на видеопоследовательностях имеют место ситуации, при которых глобальный минимум критериальной функции не соответствует истинному положению объекта. Предложена процедура анализа локальных минимумов, основанная на трёх критериях: значение критериальной функции в точке минимума, расстояние до её центра и острота вершины. На примере типовой критериальной функции, для которой происходит срыв сопровождения при использовании глобального минимума, показана эффективность предложенной процедуры. Составлена база видеопоследовательностей с размеченными треками объектов. Экспериментальные исследования показали, что применение предложенной процедуры увеличивает качество сопровождения.

### Литература

- 1. Dutta A. et al. Vision Tracking: A Survey of the State-of-the-Art // SN Computer Science. 2020. No. 1 Art. 57. 19 p.
- 2. Yao R. et al. Video object segmentation and tracking: A survey // arXiv. 2019. Vol. 1. No 1. Art. 1. 36 p.

- 3. Walia G. S., Kapoor R. Recent advances on multicue object tracking: a survey // Artificial Intelligence Review. 2016. Vol. 46. No. 1. pp. 1-39.
- 4. 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.
- 5. Алпатов Б.А., Бабаян П.В. Методы обработки и анализа изображений в бортовых системах обнаружения и сопровождения объектов // Цифровая обработка сигналов. 2006. № 2. С. 45-51.
- 6. Yao R. et al. Video object segmentation and tracking: A survey // arXiv. 2019. Vol. 1. No. 1. pp. 1-36.
- 7. Yilmaz A., Javed O., Shah M. Object tracking: A survey // ACM computing surveys (CSUR). 2006. Vol. 38. No. 4. pp. 1-45.
- 8. Алпатов Б.А., Селяев А.А. Алгоритм оценки местоположения объекта на двумерном изображении // Изв. вузов. Приборостроение. 1988. № 5. С. 3-5.
- 9. Можейко В.И., Обухова Н.А., Тимофеев Б.С., Фисенко В.Т., Фисенко Т.Ю. Телевизионные методы сопровождения объектов в сложных условиях наблюдения // Изв. вузов. Приборостроение. 2009. Т. 52. № 8. С. 20-29.
- 10. Алпатов Б.А., Бабаян П.В., Корепанов С.Е. Корреляционный алгоритм слежения за движущимися объектами при наличии близкорасположенных похожих объектов // Вестник РГРТУ. 2010. № 2 (32). С.3-8.

### Уважаемые авторы!

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.

УЛК 621.391

# МЕТОДЫ ШИРОКОПОЛОСНОЙ ПЕРЕДАЧИ ДАННЫХ НА ОСНОВЕ СИГНАЛОВ С ЧАСТОТНЫМ РАЗДЕЛЕНИЕМ КАНАЛОВ

Никишкин П.Б., н.с. кафедры TOP PГРТУ, e-mail: tor@rsreu.ru; Витязев В.В., д.т.н., заведующий кафедрой TOP PГРТУ e-mail: vityazev.v.v@rsreu.ru.

# BROADBAND DATA TRANSMISSION METHODS BASED ON SIGNALS WITH FREQUENCY-DIVISION MULTIPLEXING

### Nikishkin P.B., Vityazev V.V.

Methods of broadband data transmission with frequency division multiplexing are considered. A method for adaptive broadband data transmission based on filter bank and OFDM technologies is presented. The possibility of increasing the spectral and computational efficiency of the data transmission system is shown. The results of studying the Doppler shifts using OFDM technology are generalized.

**Key words**: broadband data transmission systems, ofdm, filter bank.

**Ключевые слова:** Широкополосные системы передачи данных, технология OFDM, банк фильтров.

#### Введение

В современных стандартах беспроводной связи используются методы передачи информации на основе сигналов со многими несущими.

С развитием аппаратных платформ стал возможным быстрый расчёт дискретного преобразования Фурье, в результате чего мультиплексирование сигналов с ортогональным частотным разделением каналов (OFDM) стало широко использоваться [1, 2].

К основным преимуществам технологии OFDM можно отнести:

- увеличение длительности символов (по отношению к методам модуляции с одной несущей), в результате чего появилась возможность противостоять межсимвольной интерференции;
- при использовании дискретного преобразования Фурье генерация OFDM сигнала вычислительно простая операция;
  - высокая спектральная эффективность;
  - простая реализация эквалайзера;
  - эффективное сочетание OFDM и MIMO [4, 9].

Для повышения эффективности в широкополосных системах со многими несущими возникают проблемы с

Рассматриваются методы широкополосной передачи данных с частотным разделением каналов. Представлен способ адаптивной широкополосной передачи данных на основе технологий банка фильтров и OFDM. Показана возможность повышения спектральной и вычислительной эффективности системы передачи данных. Обобщены результаты исследования доплеровских смещений на технологию OFDM.

межсимвольной и межканальной интерференцией. Известен целый ряд подходов для борьбы с данными проблемами. В системах связи с технологией ОFDM, для борьбы с межсимвольной интерференцией, используют защитный интервал, однако это приводит к потере спектральной эффективности. Одна из главных проблем ОFDM — это межканальная интерференция, обусловленная частотным рассеиванием и расширением, что приводит к потере ортогональности частотных каналов и появлению ошибок [3].

В варианте реализации технологии OFDM с использованием преобразования Фурье (рис. 1) вычислительные затраты можно многократно уменьшить, если количество частотных каналов кратно степени двойки.

Несмотря на простоту реализации и широкий спектр применения технологии OFDM, отмеченные выше недостатки этого популярного вида модуляции, требуют проведения дальнейших исследований эффективности работы систем высокоскоростной передачи данных [4].

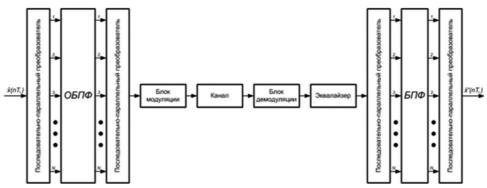


Рис. 1. Классическая схема системы передачи данных с OFDM

С другой стороны, применение технологии OFDM в таких областях как когнитивное радио, системах беспроводной передачи с многопользовательским доступом на линии «мобильный абонент — базовая станция» трудно реализуема из-за проблем с синхронизацией [5].

В этих и других приложениях применение технологии банка фильтров со многими несущими (FBMC – Filter Banks Multi-Carrier) предоставляет оптимальное решение с позиции спектральной и энергетической эффективности.

#### Технологии на основе банка фильтров

Упрощенная структура приемо-передатчика сигналов со многими несущими на основе банка фильтров представлена на рис. 2. Предполагается, что двоичный поток преобразуется в M параллельных частотных каналов  $x_k(t)$ , каждый из которых обрабатывается формирующим фильтров (ФНЧ с импульсной характеристикой h(t)) и переносится на несущую частоту, соответствующему k-му каналу [10].

Скорость параллельного цифрового потока в каждом канале уменьшается в M раз по сравнению с последовательным потоком. В результате при общей неизменной битовой скорости и неизменной занимаемой полосе частот в сигнале со многими несущими длительность импульсов в каждом канале увеличивается в M раз. В отличие от OFDM для FBMC нет необходимости использовать защитный интервал, поэтому длительность сигнала является минимально возможной [10].

Технология FBMC при наличии ограничений сигнала во временной и частотной области, обычно требует наличия большой длины ИХ фильтра (например, более

чем в 3 раза превышает длительность символа). Таким образом, требуются большие вычислительные ресурсы и память данных. Более того, технология FBMC подразумевает использование более массивной антенны для передачи сигналов, по сравнению с другими технологиями, что в свою очередь нежелательно для стандартов беспроводной связи будущих поколений [12]. Таким образом, этот недостаток ограничивает применение технологии FBMC.

Рассматриваемые выше технологии являются перспективными в своих областях применения, однако проблемы, связанные со спектральной эффективностью, остаются по-прежнему актуальными.

В результате чего был предложен метод субполосного формирования широкополосного сигнала на основе технологий OFDM и банка фильтров [4, 11].

#### Субполосная OFDM

В этом методе предполагается, что доступный частотный диапазон разбивается на субполосы и внутри каждой субполосы формируется OFDM сигнал. Групповой сигнал получается объединением всех выходов, формируемых частотными каналами (см. рис. 3).

В статье [11] рассматриваются вопросы использования соседних частотный каналов с достаточно низкой межканальной интерференцией.

Преимущество данной технологии состоит в том, что нет необходимости в синхронизации, в то время как для используемых в настоящий момент технологий необходима точная синхронизация, применение которой не предполагается в будущих радиоинтерфейсах. Одной из таких технологий, не использующих систему синхро-

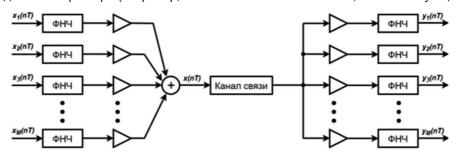


Рис. 2. Общая схема построения банка фильтров

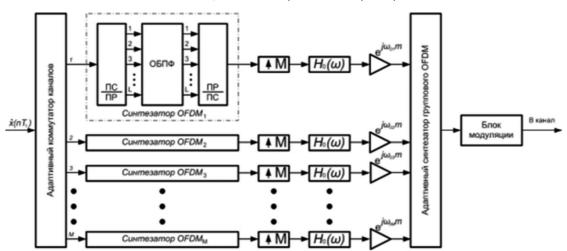


Рис. 3. Структурная схема синтезатора группового широкополосного ОFDM сигнала

		•	т		1
	OFDM	FBMC	GFDM	UFMC	RB- F-OFDM
Использование защитного интервала	+	-	-	-	-
Сложность построения фильтра	-	высокая	высокая	простая	простая
Режим модуляции	без	OQAM	без	без	без
	ограничений ООАМ		ограничений	ограничений	ограничений
Способность противостоять межсим-	низкая	высокая	высокая	высокая	высокая
вольной интерференции	КВЯСИП				
Гибкость использования спектра	-	_	-	-	+

Таблица 1. Сравнение технологий OFDM, FBMC, GFDM UFMC, F-OFDM и RB-F-OFDM

низации, является когнитивное радио. Как уже отмечалось ранее, технология FBMC может быть использована как альтернатива классическому OFDM, однако возникают проблемы, связанные со сложной реализацией эквалайзера. К достоинствам субполосного OFDM (SUB-OFDM, RB-F-OFDM — Resuorce-block filtered orthogonal frequency-division multiplexing) можно отнести обратную совместимость с существующими системами. Таким образом, приемник субполосного OFDM может принимать сигналы систем передачи информации на основе OFDM. При реализации субполосного OFDM можно эффективно использовать методы уменьшения пикфактора, для повышения энергетической эффективности системы [11].

В статье [12] оценивается производительность таких технологий как FBMC, UFMC (Universal Filtered Multi-Carrier), GFDM (Generalized Frequency-Division Multiplexing), RB-F-OFDM. Можно отметить, что системы FBMC и GFDM являются более перспективными технологиями, однако использование данных технологий требует существенной доработки существующих систем, а работа по коротким сигналам нерациональна изза низкой частотной-временной эффективности. У системы субполосного OFDM легкая обратная совместимость, поэтому развитие и исследование данной технологии остается актуальной.

Использование всех этих технологий подразумевает уменьшение внеполосного излучения и упрощение системы синхронизации. Стоит отметить, что в современных методах формирования широкополосных сигналов применяется фильтрация с различной методологией проектирования и эффективностью. Более подробное сравнение приведено в табл. 1.

В технологии GFDM можно отметить, что поднесущие расположены достаточно близко друг к другу, из-за чего нет взаимной ортогональности. Чтобы устранить межсимвольную интерференцию необходимо применение фильтров высокого порядка и добавление нулевый отсчетов. Кроме того, требуется предварительное подавление помех для ослабления их действия между поднесущими, которые могут появляться после фильтрации [12].

Технология UFMC использует длину импульсной характеристики фильтра, равную длительности защитного интервала в системе OFDM. Это позволяет уменьшать межсимвольную интерференцию, но уровень внеполосного излучения может быть неудовлетворительным. В отличие от этого, используя длину импульсной характеристики фильтра до половины длительности символа,

f-OFDM намеренно отказывается от ортогональности между последовательными символами OFDM в обмен на более низкий уровень внеполосного излучения и, таким образом, позволяет использовать минимальное количество защитных интервалов. При правильно спроектированных фильтрах (например, с ограниченным разбросом энергии) ухудшение рабочих характеристик в результате увеличения длины фильтра почти незначительно по сравнению с экономией на потреблении защитной полосы.

Технология RB-F-OFDM сочетает в себе все достоинства технологии f-OFDM и позволяет динамически подстраиваться под изменения в канале связи. Затраты на динамическое изменение фильтров превышают показатель для f-OFDM, однако их можно минимизировать при использовании частотных диапазонов одинаковой длительности.

В целом технология RB-F-OFDM выглядит как наиболее многообещающий претендент на базовый сигнал для 5G, обеспечивая известные преимущества OFDM:

- гибкое частотное мультиплексирование;
- простое выравнивание каналов;
- простая комбинация с многоантенной передачей, но и новые достоинства:
- услуги, адаптированные к различным потребностям;
  - эффективное использование спектра;
  - низкое значение внеполосного излучения;
  - доступная вычислительная сложность;
  - возможность включения сигналов других форм;
  - обратная и прямая совместимость.

В статьях [11, 12] технология субполосного OFDM (RB-F-OFDM) рассматривается с точки зрения спектральной эффективности и вычислительных затрат, проводятся исследования работы в условиях канала с белым шумом, но при этом не учитывается работа в условиях воздействия эффекта Доплера.

В статье [7] проведено исследование работы технологии в канале с доплеровскими искажениями. Если обобщить результаты, приведённые в статьях [7, 8], то можно сделать вывод, что доплеровские искажения приводят к повороту сигнально-кодового созвездия на некоторый угол (см. рис. 4) и значительных искажений в принимаемый сигнал не вносят.

Расстройка временной синхронизации приводит к вращению точек на сигнально-кодовом созвездие (см. рис. 5). Такое движение точек можно объяснить изменением частоты дискретизации и получением амплитуд

точек в моменты времени, не соответствующие максимальному значению.

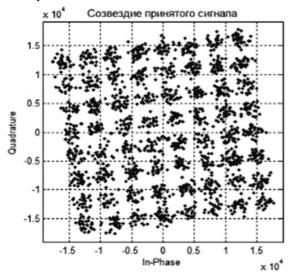


Рис. 4. Сигнально-кодовое созвездие принятого сигнала при наличии доплеровского сдвига в 200 Гц

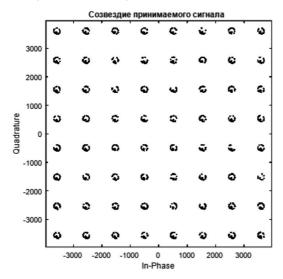


Рис. 5. Сигнально-кодовое созвездие принятого сигнала при наличии расстройки временной синхронизации

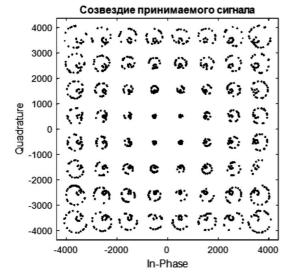


Рис. 6. Сигнально-кодовое созвездие принятого сигнала при наличии многолучевости

Как показали исследования, многолучевость одинаково влияет на работу технологий OFDM и SUB-OFDM. При условии наличия многолучевости выигрыш при использовании технологии SUB-OFDM получить сложно.

Из представленных рисунков выше можно сделать выводы, что технология SUB-OFDM работает так же, как и технология OFDM при моделировании простых различных мешающих воздействий. Требуется дальнейшее исследование работоспособности и эффективности технологии при более сложных мешающих воздействиях. Применение технологии SUB-OFDM позволяет в значительной степени подавить внеполосное излучение, межканальные интерференции и повысить эффективность передачи данных, за счет значительного уменьшения защитного интервала. Применение данной технологии позволяет в разрабатываемой системе гибко использовать частотный ресурс. Благодаря набору цифровых полосовых фильтров возможна передача в несмежных субполосах.

Дальнейшие исследования эффективности технологии субполосного OFDM следует проводить на таких искажениях и рассогласованиях в системах с OFDM как, фазовый шум несущей, разбалансировка квадратурных составляющих, и проблемы, вызываемые пик-фактором.

Статья подготовлена при поддержке НИИ «Фотон», Рязанский государственный радиотехнический университет имени В.Ф. Уткина.

#### Заключение

Произведён обзор известных технологий, используемых для передачи данных со многими несущими. Обобщена информация исследования технологии субполосного OFDM в условиях работы с доплеровскими эффектами. Использование данной технологии позволяет значительно уменьшить использование защитного интервала, и тем самым повысить спектральную эффективность. При применении в системах передачи данных наблюдается уменьшение затрат на реализацию эквалайзера и систем синхронизации. Предполагается, что данную технологию нужно использовать в разрабатываемых в настоящее время стандартах беспроводной связи нового поколения (например, 5G), так как в них требуется обеспечение более высокой скорости передачи данных и эффективного использования частотного ресурса. Технология субполосного OFDM обратно совместима с классическим OFDM, поэтому приемником субполосных OFDM сигналов можно принимать и демодулировать сигналы текущих беспроводных стандартов связи. Внедрение субполосного OFDM в текущие стандарты связи требует модернизации только передатчика и приемника OFDM сигналов, в результате будет наблюдаться увеличение скорость передачи данных. Определён путь дальнейших исследований перспективного метода субполосной передачи данных.

### Литература

1. Витязев В.В. Цифровая обработка сигналов: ретроспектива и современное состояние // Электросвязь. – 1997. – № 6.

- 2. Шойерманн X., Геклер X. Систематизированный обзор цифровых методов преобразования вида уплотнения каналов // ТИИЭР. 1981. Т. 69. № 11. С. 52-84.
- 3. Behrouz Farhang-Boroujeny. OFDM Versus Filter Bank Multicarrier // IEEE Signal Processing Magazine, 2011, vol. 28, no. 3, pp. 92-112.
- 4. Витязев В.В., Никишкин П.Б. Банки фильтров и OFDM в системах широкополосной передачи данных со многими несущими. // Научно-технический журнал «ЦОС». 2015. №4. С. 30-34.
- 5. Витязев В.В., Никишкин П.Б. Метод анализа/синтеза сигналов в системах передачи данных с частотным уплотнением каналов. // ЭЛЕКТРОСВЯЗЬ. 2014. № 12. С. 4-9.
- 6. Витязев В.В., Никишкин П.Б.Исследование эффектов Доплера на OFDM и SUB-OFDM сигналы // 1-я Всероссийская конференция «Современные технологии обработки сигналов», Москва, Россия, доклады конференции.
- 7. Витязев В.В., Никишкин П.Б. Сравнение эффективности технологий OFDM и SUB-OFSM при различных мешающих воздействиях в канале связи. // 21-я Между-

- народная конференция «Цифровая обработка сигналов и ее применение DSPA-2019», Москва, Россия, доклады. 2019. Книга 1. С. 6-10.
- 8. Майков Д.Ю., Вершинин А.С. Влияние эффектов Доплера на OFDM сигнал // Молодой ученый. 2014. №21. С. 175-179. URL https://moluch.ru/archive/80/14271/ (дата обращения: 18.11.2019).
- 9. Бакулин М.Г., Крейнделин В.Б., Шлома А.М., Шумов А.П. Технология ОFDM: Учебное пособие для вузов / М.: Горячая линия-Телеком, 2017. 352 с. ISBN 978-5-9912-0549-8
- 10. Витязев В.В., Овинников А.А. Методы анализа/ синтеза сигналов в системах беспроводной связи со многими несущими // ЭЛЕКТРОСВЯЗЬ. 2013. № 9. С. 28-32.
- 11. Jialing Li and Erdem Bala and Rui Yang Resource block Filtered-OFDM for future spectrally agile and power efficient systems. Physical Communication, 2014/11, pp. 36-55, doi: 10.1016/j.phycom.2013.10.003
- 12. Van Eeckhaute M., Bourdoux A., De Doncker P. et al. Performance of emerging multi-carrier waveforms for 5G asynchronous communications. J Wireless Com Network 2017, 29 (2017) doi:10.1186/s13638-017-0812-8.

# НОВЫЕ КНИГИ





Основы теории цифровой обработки сигналов: Учебное пособие. Издание пятое, исправленное и дополненное. М: ТЕХНОСФЕРА, 2019. — 550 с.: ил.

Учебное пособие предназначено для студентов вузов, обучающихся по направлениям подготовки бакалавров и магистров «Прикладная математика», «Информатика и вычислительная техника», и направлено на изучение теоретических основ цифровой обработки сигналов (ЦОС). Помимо базовых разделов ЦОС, связанных с цифровым представлением сигналов, анализом и синтезом цифровых фильтров, в пособии рассматриваются вопросы многоскоростной обработки сигналов, адаптивной фильтрации, дискретного спектрального анализа. Значительное внимание уделено таким специальным разделам ЦОС, как эффективное представление сигналов (сжатие данных), теория и приложения дискретных вейвлет-преобразований.

Пособие также может быть рекомендовано в качестве дополнительного учебного материала для других инженерных направлений подготовки, связанных с ЦОС (радиотехника, электроника, телекоммуникации и связь, автоматическое управление и др.). При этом базовая часть материала первых глав представляет собой основу для вводного курса (уровень бакалавриата), а в полном объеме учебное пособие ориентировано на более углубленное изучение теории ЦОС в магистратуре.



Проектирование функциональных блоков встраиваемых систем на FPGA-M.: Изд-во «Горячая линия-Телеком», 2020 г. -348 с.: ил.

Рассмотрены основы проектирования функциональных блоков, которые наиболее часто встречаются во встраиваемых системах, на программируемых пользователем вентильных матрицах FPGA (field programmable gate array – FPGA). Рассмотрено проектирование на FPGA путем описания на языке Verilog и с помощью IP-ядер блоков памяти различного типа: одно-портовой, двух-портовой RAM, ROM, FIFO, LIFO, а также сдвиговых регистров в блоках встроенной памяти. Представлено несколько методик проектирования устройств управления: в виде микропрограммного автомата (МПА) по граф-схеме алгоритма (ГСА), на основе блок-схем автоматов (ASM), а также на основе блок-схем автоматов с трактом обработки данных (FSMD). Приведена методика проектирования на FPGA одно-тактового процессора, а также три методики проектирования много тактовых процессоров. Особое внимание уделено отладке процессора PIC и оценке его производительности, приводятся рекомендации

по увеличению производительности процессоров. Представлена общая методология проектирования цифровых фильтров, рассмотрен пример разработки цифрового фильтра в системе MATLAB, моделирования фильтра в системе ModelSim и реализация фильтра на FPGA в системе Quartus. Рассмотрены вопросы проектирования подсистем синхронизации встраиваемых систем на FPGA. Описаны принципы функционирования блоков фазовой автоподстройки частоты PLL, особенности архитектуры и функционирования блоков PLL в FPGA, а также способы конфигурирования блоков PLL в проектах на FPGA.

Для специалистов в области разработки встраиваемых систем, научных работников, аспирантов.



УЛК 621.372

# РОЛЬ И МЕСТО ОПЕРАЦИИ ДОПОЛНЕНИЯ НУЛЯМИ В ТЕОРИИ ДВУМЕРНОЙ ФУРЬЕ-ОБРАБОТКИ СИГНАЛОВ

Пономарев А.В., к.э.н., доцент Ижевского государственного технического университета имени М.Т. Калашникова, e-mail: ponva@ mail.ru;

Пономарева О.В., д.т.н., профессор, Ижевского государственного технического университета имени М.Т. Калашникова, е-mail: ponva@ mail.ru.

# THE ROLE OF ZERO PADDING IN THE THEORY OF TWO-DIMENSIONAL FOURIER SIGNAL PROCESSING

#### Ponomarev A.V., Ponomareva O.V.

The transition to two-dimensional Fourier processing requires rethinking many concepts and definitions of one-dimensional digital Fourier processing of signals. For example, the operation of appending zeros to the original signal in one-dimensional Fourier signal processing is an effective method for eliminating aliasing effects, detailing the spectrum estimate of finite discrete one-dimensional signal. In two-dimensional Fourier signal processing, the corresponding operation is also effective, but requires rethinking. The paper presents the systems analysis of theoretical foundations of discrete two-dimensional signal processing based on Fourier transform. Evaluation of the efficiency of the zero-padding operation in two-dimensional signal processing is carried out.

The concept of discrete-spatial Fourier transform is introduced. It is shown that the discrete-spatial Fourier transform is defined as a two-dimensional z-transform. This transformation is computed in z-space on the unit sphere.

An approximation of the discrete - spatial Fourier transform is considered. The approximation of the discrete-spatial Fourier transform is based on a two-dimensional discrete Fourier transform of a zero-padded signal.

A systems analysis of the postulates of the theory of discrete two-dimensional signal processing based on Fourier transform is given. Methods and algorithms for obtaining a two-dimensional linear convolution using cyclic convolution are presented. Methods and algorithms for obtaining a two-dimensional linear correlation function based on a cyclic correlation function are presented. The results of numerical simulation are presented, which confirm the obtained theoretical results.

**Key words:** two-dimensional signal, reference domain, two-dimensional discrete Fourier transform, two-dimensional discrete-spatial Fourier transform, two-dimensional convolution, two-dimensional correlation function, zero padding.

**Ключевые слова:** двумерный сигнал, опорная область, двумерное дискретное преобразование Фурье, двумерное дискретно – пространственное преобразование Фурье, двумерная свертка, двумерная корреляционная функция, операция дополнения нулями.

### Введение

Переход от одномерной (1-D) к двумерной (2-D) Фурье – обработке требует переосмысления многих понятий и определений 1-D цифровой обработки сигналов (ЦОС) [1-30]. Это с одной стороны объясняется существенными различиями в теориях 1-D и 2-D линейных систем. С другой стороны, необходимость переосмысления связана с тем, что переход от 1-D к 2-D обработке в базисах Фурье сигналов является не только количественным, но и в существенной степени качественным переходом. Среди понятий и определений 1-D ЦОС, которые требуют переформулирования при переходе к 2-D ЦОС такие как: свертка, корреляция, эффекты наложения, частокола, утечки. Операция дополнения исходного сигнала нулевыми отсчетами в 1-D Фурье - обработке, которая является эффективным и результативным методом детализации оценки спектра финитных, дискретных, одномерных сигналов<sup>1</sup>, также не является исключением.

Целью данной работы является системный анализ

Проведен системный анализ теоретических основ дискретной двумерной обработки сигналов на базе преобразований Фурье. Получена оценка эффективности операции дополнения нулями в двумерной обработке сигналов.

Введено понятие дискретно-пространственного преобразования Фурье. Показано, что дискретно-пространственное преобразование Фурье определяется как двумерное z-преобразование. Данное преобразование вычисляется в z-пространстве на единичной сфере.

Рассмотрена аппроксимация дискретно-пространственного преобразования Фурье, основанная на двумерном дискретном преобразовании Фурье дополненного нулями сигнала.

Дан системный анализ постулатов теории дискретной двумерной обработки сигналов на базе преобразований Фурье. Приведены методы и алгоритмы получения двумерной линейной свертки и двумерной линейной корреляции с помощью циклической свертки на основе циклической свертки и циклической корреляционной функции. Приведены результаты численного моделирования, подтверждающие полученные теоретические результаты.

теоретических основ дискретной двумерной обработки сигналов на базе преобразований Фурье, оценка роли и места в данной теории операции дополнения нулями.

<sup>&</sup>lt;sup>1</sup> **Финимным дискретным сигналом** называют сигнал x(n), определенный при целочисленных значениях n, при  $-\infty \le n \le +\infty$  и отличный от нуля на интервале длительностью в N отсчетов.

# Анализ теоретических основ дискретной двумерной обработки сигналов на базе преобразований Фурье

Финитной дискретной 2-D последовательности  $x(n_1,n_2)\;;\;n_1=\overline{0,N_1-1},\;n_2=\overline{0,N_2-1}\;; \tag{1}$  с прямоугольной опорной областью  $SA_{N_1\times N_2}$  :

$$SA_{N_1 \times N_2} := \{ (n_1, n_2) : n_1 = \overline{0, N_1 - 1}, n_2 = \overline{0, N_2 - 1} \};$$
 (2)

можно поставить в соответствие двумерную периодическую последовательность  $x_p(n_1,n_2)$  с фундаментальным периодом  $SA_{N_1 N_2 N_2}^{2}$ :

$$x_p(n_1, n_2) = x(n_1 + l_1 \cdot N_1, n_2 + l_2 \cdot N_2);$$
 (3)  $l_1, l_2$  — целые числа.

В силу того, что 2-D ДПФ  $S_{N_1,N_2}(k_1,k_2)$  дискретной 2-D последовательности  $x(n_1,n_2)$  определяется соотношением:

$$S_{N_1,N_2}(k_1,k_2) = \frac{1}{N_1 \cdot N_2} \sum_{n=0}^{N_1-1} \sum_{n_1=0}^{N_2-1} x(n_1,n_2) W_{N_1}^{k_1 n_1} \cdot W_{N_2}^{k_2 n_2};$$
 (4)

где 
$$W_{N_1}^{k_1n_1}=\exp\biggl(-jrac{2\pi}{N_1}k_1n_1\biggr); \quad W_{N_2}^{k_2n_2}=\exp\biggl(-jrac{2\pi}{N_2}k_2n_2\biggr)$$
 ;

то спектру  $S_{N_1,N_2}(k_1,k_2)$  можно поставить в соответствии 2-D последовательность  $S_{p,\;N_1,N_2}(k_1,k_2)$  с фундаментальный периодом  $SA_{N,\times N_2}$  :

$$S_{p, N_1, N_2}(k_1, k_2) = S_{N_1, N_2}(k_1 + l_1 \cdot N_1, k_2 + l_2 \cdot N_2);$$
 (5)  $l_1, l_2$  – целые числа.

В теории цифровой Фурье – обработки 1-D дискретных сигналов широкое применение нашло *дискретновременное преобразование Фурье* (ДВПФ). ДВПФ – это z-преобразование, вычисленное в z-плоскости на единичной окружности.

В двумерном случае можно ввести аналогичное ДВПФ преобразование — *дискретно-пространственное преобразование Фурье* (ДППФ). ДППФ — это двумерное z-преобразование, вычисленное в z-пространстве на единичной сфере:

$$X_{\text{ДПП}\Phi}(f_1, f_2) = \sum_{m = -\infty}^{\infty} \sum_{n = -\infty}^{\infty} x(m, n) \cdot z_1^{-m} z_2^{-n} \Big|_{z_2 = \exp(+j2\pi f_2)}^{z_1 = \exp(+j2\pi f_2)}; \quad (6)$$

где  $f_1$  и  $f_2$  непрерывные пространственные частоты.

Сравнивая соотношения (4) и (6) несложно установить, что двумерное ДПФ сигнала  $x(n_1,n_2)$  равно двумерному ДППФ, вычисленному на единичной сфере в точках  $\exp(-j2\pi\,k_1\,/\,N_1)$  и  $\exp(-j2\pi\,k_2\,/\,N_2)$ .

Отметим, что ДВПФ и ДППФ представляют собой соответственно непрерывные преобразования дискретных 1-D и 2-D последовательностей и являются чисто теоретическими понятиями одномерной теории и двумерной теории цифровой Фурье – обработки.

ДППФ финитного дискретного сигнала (1) с фундаментальным периодом  $SA_{N_1 \times N_2}$  может быть аппроксимировано с помощью операции дополнения нулями путем вычисления 2-D ДПФ сигнала  $x_1(m,n)$  с фундаментальным периодом  $SA_{L_1 \times L_2}$  размерностью  $L_1 \times L_2$ :

$$x_{1}(m,n) = \begin{cases} x(m,n); & m = \overline{0, N_{1} - 1}; & n = \overline{0, N_{2} - 1} \\ 0; & m = \overline{N_{1}, L_{1} - 1}; & \overline{n} = N_{2}, L_{2} - 1 \end{cases}$$
 (7)

Анализ теоретических основ дискретной двумерной обработки сигналов на базе преобразований Фурье позволяет сделать следующие важные выводы.

Двумерное дискретное преобразование Фурье следует понимать как двумерный дискретный ряд Фурье, из свойств которого во многом вытекают свойства 2-D ДПФ.

Свойства двумерного ДПФ (как и свойства одномерного ДПФ) являются математически точными, и не могут трактоваться как аппроксимация свойств непрерывного преобразования Фурье (НПФ). В то же время одномерное и двумерное ДПФ конечной последовательности могут рассматриваться как аппроксимации соответствующих НПФ.

Теория двумерной обработки дискретных сигналов на основе двумерного дискретного преобразования Фурье является теорией двумерной Фурье - обработки финитных двумерных дискретных сигналов.

Проведенный анализ теоретических основ дискретной двумерной обработки сигналов позволил сформулировать следующие основополагающие и взаимоувязанные ее аксиоматические положения (постулаты):

Определение двумерных дискретных сигналов на конечной опорной области – фундаментальном периоде, задаваемом вертикальным и горизонтальным периодами.

Определение сдвига двумерного дискретного сигнала как циклического сдвига в виде циклической перестановки его отсчетов на конечной опорной области. Принцип двумерной циклической перестановки отсчетов сигнала на конечной опорной области заключается в следующем: если справа или сверху отсчет 2-D сигнала «вышел», то этот отсчет 2-D сигнала соответственно слева или снизу «зашел». Циклическая перестановка пространственных и пространственно-частотных отсчетов двумерного дискретного сигнала, может быть представлена периодическим продолжением его отсчетов (с фундаментальным периодом) за пределами опорной области. Циклическая перестановка пространственных и пространственно-частотных отсчетов двумерного дискретного сигнала может быть также предпространственными пространственноставлена И частотными отсчетами, расположенными на единичной сфере в соответствующей области.

Определение полной базисной системы 2-D ДПФ в виде системы двумерных дискретных экспоненциальных функций.

Отметим, что конечность, периодичность и дискретность последовательностей в пространственной или пространственно-частотной области приводит к появлению новых свойств как 1-D ДПФ так и 2-D ДПФ, которые отсутствуют в непрерывных преобразованиях Фурье. Конечность, периодичность и дискретность последовательностей в пространственной и простран-

 $<sup>^{2}</sup>$  **Фундаментальный период**  $SA_{N_{1}\times N_{2}}$  определяется соответственно вертикальным и горизонтальным периодами  $N_{1}$  и  $N_{2}$ .

ственно-частотной области приводит также к видоизменению некоторых понятий Фурье — обработки непрерывных сигналов (например, видоизменению таких понятий как сдвиг, свертка и корреляция). В непрерывном случае понятие линейного сдвига двумерного сигнала  $x(t_1,t_2)$  в пространственной области (как и одномерного во временной области) особого пояснения не требует, поскольку соответствует интуитивному пониманию сдвига. Введем в рассмотрение еще одно математическое определение циклического сдвига дискретного сигнала  $x(n_1,n_2)$  при изменении переменных  $n_1 \geq N_1$  и  $n_2 \geq N_2$ :

$$x(n_1, n_2) = x(((n_1))_{\text{mod } N_1}, ((n_2))_{\text{mod}_{N_2}});$$
  

$$(n_1, n_2) \in SA_{N_1 \times N_2};$$
(8)

где  $((n))_{\text{mod }N}$  – символ взятия числа по модулю  $N^3$ .

Если 2-D ДПФ сигнала  $x(n_1,n_2)\,$  равно –  $S_{\scriptscriptstyle N_1,N_2}(k_1,k_2)$  :

$$x(n_1, n_2) \stackrel{F}{\longleftrightarrow} S_{N_1, N_2}(k_1, k_2); \tag{9}$$

то 2-D ДПФ сигнала  $x(((n_1-p_1))_{\mod N_1}),((n_2-p_2))_{\mod N_2}$  умножается на двумерную экспоненту:

$$x(((n_1 - p_1))_{\text{mod } N_1}), ((n_2 - p_2))_{\text{mod } N_2} \overset{F}{\longleftrightarrow} W_{N_1}^{p_1 k_1} \cdot W_{N_2}^{p_2 k_2}$$

$$S_{N_1, N_2}(k_1, k_2) . \tag{10}$$

Известно, что если:

$$x(n_1, n_2) \xrightarrow{F} S_{N_1, N_2}(k_1, k_2), \text{ if } y(n_1, n_2) \xrightarrow{F} Y_{N_1, N_2}(k_1, k_2)$$

TO:

$$x(n_1, n_2) * y(n_1, n_2) \xrightarrow{F}$$

$$S_{N_1, N_2}(k_1, k_2) \cdot Y_{N_1, N_2}(k_1, k_2);$$
(11)

где \* - символ свертки.

Выражение (11) задает так называемую циклическую свертку  $h(n_1,n_2)$  двух 2-D сигналов  $x(n_1,\ n_2)$  и  $y(n_1,\ n_2)$  в пространственной области:

$$h(n_1, n_2) =$$

$$= \sum_{p_1}^{N_1-1} \sum_{p_2}^{N_2} x(p_1, p_2) \cdot y(((n_1 - p_1))_{\text{mod } N_1}, ((n_2 - p_2))_{\text{mod } N_2}); \quad (12)$$

$$(n_1, n_2) \in SA_{N_1 \times N_2};$$

ипи:

$$h(n_1, n_2) =$$

$$=\sum_{p_1}^{N_1-1}\sum_{p_2}^{N_2}y(p_1,p_2)\cdot x(((n_1-p_1))_{\mathrm{mod}\,N_1},((n_2-p_2))_{\mathrm{mod}\,N_2}); (13)$$

$$(n_1, n_2) \in SA_{N_1 \times N_2}$$
.

Циклическая 2-D корреляционная функция определяется следующими соотношениями:

$$c(n, n_2) = x(n_1, n_2) * x(n_1, n_2) \xrightarrow{F} \left| S_{N_1, N_2}(k_1, k_2) \right|^2;$$
 (14)

$$c(n_1,n_2)=$$

$$=\sum_{p_1}^{N_1-1}\sum_{p_2}^{N_2}x(p_1,p_2)\cdot x(((n_1-p_1))_{\mathrm{mod}\,N_1},((n_2-p_2))_{\mathrm{mod}\,N_2}); \quad \text{(15)}$$
 $(n_1,n_2)\in SA_{N_1\times N_2};$ 
или:

$$c(n_1, n_2) =$$

$$= \sum_{p_1}^{N_1-1} \sum_{p_2}^{N_2} x(p_1, p_2) \cdot x(((n_1-p_1))_{\text{mod } N_1}, ((n_2-p_2))_{\text{mod } N_2}); \quad (16)$$

$$(n_1, n_2) \in SA_{N_1 \times N_2}$$
.

В то же время при разработке и моделировании изопланатических систем, Фурье — обработке 2-D сигналов, необходима линейная 2-D свертка (линейная 2-D корреляционная функция).

Метод, позволяющий с помощью 2-D циклической свертки (2-D циклической корреляционной функции) получить линейную 2-D свертку (2-D корреляционную функцию), как отмечалось выше, заключается в устранении наложения 2-D сигналов в корреляционной области путем соответствующего дополнения сворачиваемых дискретных сигналов нулевыми отсчетами в их опорных областях.

Если задана опорная область  $SA_{V_1 \times V_2}$  сигнала  $x(n_1,n_2)$  и опорная область  $SA_{Q_1 \times Q_2}$  сигнала  $y(n_1,n_2)$ , то размер опорной области (фундаментального периода) для получения линейной свертки  $h_{\scriptscriptstyle \it RMH}(n_1,n_2)$  должен быть:

$$SA_{(V_1+Q_1)\times(V_2+Q_2)}$$
; (17) где  $n_1 = \overline{0, (V_1+Q_1-1)}$ ;  $n_2 = \overline{0, (V_2+Q_2-1)}$ .

А размер опорной области для получения линейной корреляции сигнала  $x(n_1,n_2)$  должен быть:

$$SA_{2V_1 \times 2V_2}$$
; где  $n_1 = \overline{0,(2V_1 - 1)}$ ;  $n_2 = \overline{0,(2V_2 - 1)}$ . (18)

Следовательно, алгоритм получения 2-D линейной свертки на основе 2-D циклической свертки состоит из следующих операций.

# Алгоритм получения 2-D линейной свертки на основе 2-D циклической свертки

1. Дополнить 2-D сигналы  $x(n_1,n_2)$  и  $y(n_1,n_2)$  соответственно  $Q_1$ ,  $Q_2$  и  $V_1$ ,  $V_2$  нулевыми отсчетами, задав, таким образом, новые 2-D сигналы  $x_0(n_1,n_2)$  и  $y_0(n_1,n_2)$  с горизонтальными  $N_2$  и вертикальными  $N_1$  периодами согласно соотношениям:

$$N_1 \ge (V_1 + Q_1 - 1); \ N_2 \ge (V_2 + Q_2 - 1).$$
 (19)

2. Выполнить 2-D ДПФ двумерных сигналов  $x_0(n_1,n_2)$  и  $y_0(n_1,n_2)$  :

$$x_{0}(n_{1}, n_{2}) \xrightarrow{F} X_{0, N_{1}, N_{2}}(k_{1}, k_{2});$$

$$y_{0}(n_{1}, n_{2}) \xrightarrow{F} Y_{0, N_{1}, N_{2}}(k_{1}, k_{2}).$$
(20)

3. Выполнить 2-D ОДПФ произведения

$$X_{0,N_1,N_2}(k_1, k_2) \cdot Y_{0,N_1,N_2}(k_1, k_2)$$
 (21)

Алгоритм получения линейной 2-D корреляционной функции на основе циклической 2-D корреляционной

<sup>&</sup>lt;sup>3</sup> **Каждое целое число** k можно представить в виде суммы  $k = entire(k/m) + ((k))_{mod m}$ , где entire(ahr. floor) — символ взятия целой части числа k, m — положительное число, называемое модулем,  $((k))_{mod m}$  — остаток.

функции аналогичен предыдущему алгоритму. Если задана опорная область  $SA_{N_1 \times N_2}$  сигнала  $x(n_1,n_2)$ , то размер опорной области (фундаментального периода) для получения линейной корреляционной функции  $c_{\mathit{nun}}(n_1,n_2)$  должен быть:

$$SA_{2N_1 \times 2N_2}$$
; где  $n_1 = \overline{0, 2N_1 - 1}$ ;  $n_2 = \overline{0, 2N_2 - 1}$ . (22)

Для получения 2-D линейной корреляционной функции на основе 2-D циклической корреляционной функции необходимо выполнить следующие операции.

# Алгоритм получения 2-D корреляционной функции на основе 2-D циклической корреляционной функции

1. Дополнить 2-D сигнал  $x(n_1,n_2)$  соответственно  $N_1,\ N_2$  нулевыми отсчетами, создав, таким образом, новый 2-D сигнал  $x_0(n_1,n_2)$  с горизонтальным  $2N_2$  и вертикальным  $2N_1$  периодами.

2. Выполнить 2-D ДПФ двумерного сигнала  $x_0(n_1,n_2)$  :

$$x_0(n_1, n_2) \xrightarrow{F} X_{0, N_1, N_2}(k_1, k_2);$$
 (23)

3. Выполнить 2-D обратное ДПФ произведения 
$$X_{0,\,N_1,\,N_2}(k_1,\,k_2)\cdot X_{0N_1,\,N_2}(k_1,\,k_2)$$
 . (24)

Проведем сравнение линейной и циклической корреляционных функций на примере дискретного, финитного, единичного 2-D сигнала  $x\ (n_1,n_2)$ :

$$x(n_1,n_2) = \begin{cases} 1; \ \forall \ n_1 = \overline{0,N_1-1}; \ \forall \ n_2 = \overline{0,N_2-1}; \\ 0; \ \forall \ n_1 \geq N_1; \ \forall \ n_2 \geq N_2 \end{cases};$$
 при  $N_1 = N_2 = 4$ ;  $x(n_1,n_2) = \begin{bmatrix} 0 & 1 & 2 & 3 \\ 1 & 1 & 1 & 1 \\ 2 & 1 & 1 & 1 \\ 3 & n_1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}^{n_2}$ . (25)

Циклическая 2-D корреляционная функция дискретного, финитного, единичного 2-D сигнала (25), вычисленная согласно формуле (14), в графической форме приведена на рис. 1.

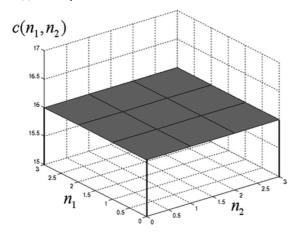


Рис. 1. Циклическая 2-D корреляционная функция дискретного, финитного, единичного 2-D сигнала (25)

Применим выше рассмотренный алгоритм получения линейной 2-D корреляционной функции на основе циклической 2-D корреляционной функции. Дополним дискретный, финитный, единичный 2-D сигнал (25) нулевыми отсчетами:

Линейная 2-D корреляционная функция сигнала  $x_0(n_1,n_2)$  в матричной форме, вычисленная согласно формуле (24), равна:

$$c_0(n_1, n_2) = \begin{bmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & n_2 \\ 1 & 2 & 3 & 4 & 3 & 2 & 1 & 0 \\ 2 & 4 & 6 & 8 & 6 & 4 & 2 & 0 \\ 3 & 6 & 9 & 12 & 9 & 6 & 3 & 0 \\ 4 & 8 & 12 & 16 & 12 & 8 & 4 & 0 \\ 3 & 6 & 9 & 12 & 9 & 6 & 3 & 0 \\ 5 & 4 & 6 & 8 & 6 & 4 & 2 & 0 \\ 6 & 7 & 2 & 3 & 4 & 3 & 2 & 1 & 0 \\ 7 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

$$n_1$$

$$(27)$$

Линейная 2-D корреляционная функция сигнала  $x_0(n_1,n_2)$  в графической форме приведена на рис. 2:

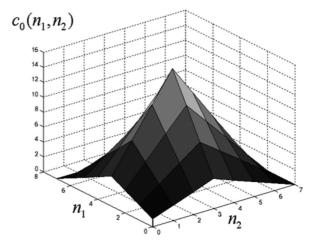


Рис. 2. Линейная 2-D корреляционная функция дискретного, финитного, единичного 2-D сигнала (25)

В заключение отметим, что в двумерном случае возможны три варианта детализации 2-D спектра дискретного, финитного 2-D сигнала – по числу вариантов дополнения опорной области нулевыми отсчетами. Если задан в матричной форме 2-D сигнал  $X_{N,\times N}$ , и задан в матричной

форме 2-D сигнал с нулевыми отсчетами  $O_{\scriptscriptstyle N,\times N_{\scriptscriptstyle 1}}$  :

$$O_{N_{1}\times N_{2}} = \begin{bmatrix} 0 & 1 & 2 & \dots & (N_{2}-1) & n \\ 0 & 0 & 0 & \dots & 0 \\ 1 & 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 0 \end{bmatrix};$$
(29)

где o(m,n) = 0:  $\forall m = \overline{0, N_1 - 1}; n = \overline{0, N_2 - 1}$ .

Тогда возможно три варианта расширения опорной области 2-D сигнала  $X_{\scriptscriptstyle N_i \times N_i}$  нулевыми отсчетами:

**Вариант 1.** Дополнение вертикального периода 2-D сигнала  $X_{N_1 \times N_2}$   $(r_1 - 1)$  нулевыми матрицами приводит к блочной матрице:

**Вариант 2.** Дополнение горизонтального периода 2-D сигнала  $X_{N_1 \times N_2}$   $(r_2-1)$  нулевыми матрицами приводит к блочной матрице:

$$0 \qquad 1 \qquad . \qquad . \qquad (r_2 - 1) \\ X_{N_1 \times N_2 \cdot r_2} = [ \quad X_{N_1 \times N_2} \quad O_{N_1 \times N_2} \quad . \quad . \quad O_{N_1 \times N_2} \quad ] \ .$$
 (31)

**Вариант 3.** Дополнение  $(r_1-1)$  нулевыми матрицами вертикального и  $(r_2-1)$  горизонтального периодов 2-D сигнала  $X_{N_1 \times N_2}$  приводит к блочной матрице:

$$X_{N_{1} \cdot r_{1} \times N_{2} \cdot r_{2}} = 0 \qquad 1 \qquad (r_{2} - 1)$$

$$0 \qquad 1 \qquad (r_{2} - 1)$$

$$1 \qquad O_{N_{1} \times N_{2}} \qquad O_{N_{1} \times N_{2}} \qquad O_{N_{1} \times N_{2}}$$

$$O_{N_{1} \times N_{2}} \qquad O_{N_{1} \times N_{2}} \qquad O_{N_{1} \times N_{2}}$$

$$\vdots \qquad \vdots \qquad \vdots \qquad \vdots$$

$$O_{N_{1} \times N_{2}} \qquad O_{N_{2} \times N_{3}} \qquad O_{N_{2} \times N_{3}} \qquad O_{N_{2} \times N_{3}}$$

$$(32)$$

Двумя существенными недостатками методов, основанных на операции дополнения сигнала в простран-

ственной области нулевыми отсчетами, являются необходимость существенного увеличения объема памяти и необходимость проведения значительного числа непроизводительных вычислений с нулевыми отсчетами.

#### Заключение

Проведенный анализ роли и места операции дополнения нулями в теории двумерной Фурье-обработки сигналов позволяет сделать следующие выводы.

- 1. Системный анализ теоретических основ дискретной двумерной обработки сигналов на базе преобразований Фурье позволяет утверждать, что свойства двумерного дискретного преобразования Фурье являются математически точными и не основаны на аппроксимации свойств непрерывного преобразования Фурье.
- 2. Теория двумерной обработки дискретных сигналов на основе двумерного дискретного преобразования Фурье является теорией двумерной Фурье обработки финитных двумерных дискретных сигналов, основанной на трех основополагающих и взаимоувязанных аксиоматических положениях: определение конечной опорной области, определении сдвига двумерного дискретного сигнала и определении полной базисной системы в виде системы двумерных дискретных экспоненциальных функций.
- 3. Операция дополнения двумерного сигнала нулевыми отсчетами играет важнейшую роль в теории дискретной двумерной обработки сигналов на базе преобразований Фурье, поскольку позволяет разрешить многие проблемы Фурье обработки и, как следствие, расширить спектр предметных областей приложения данной теории.
- 4. Операция дополнения двумерного сигнала нулевыми отсчетами является методом аппроксимации двумерного дискретно пространственного преобразования Фурье, введенного в данной работе.
- 5. Операция дополнения двумерного сигнала нулевыми отсчетами в два раза позволяет находить с помощью циклической свертки и циклической корреляционной функции соответственно линейную свертку и линейную корреляционную функцию.
- 6. Дополнение двумерного сигнала нулевыми отсчетами более чем в два раза позволяет значительно детальнее анализировать 2-D энергетический спектр дискретного сигнала, опираясь на двумерный вариант теоремы Винера-Хинчина и двумерное дискретное преобразование Фурье.
- 7. Необходим поиск эффективных и результативных методов и алгоритмов Фурье обработки дискретных, двумерных, финитных, дополненных нулевыми отсчетами.

#### Литература

- 1. Рабинер Л., Гоулд Б. Теория и применение цифровой обработки сигналов: Перевод с англ. М.: Мир. 1978. 839 с.
- 2. Марпл-мл. С.Л. Цифровой спектральный анализ и его приложения: Перевод с англ. М.: Мир, 1990. 584 с.
- 3. Gonzalez R.C., Woods R.E. Digital Image Processing, 4<sup>th</sup> Ed. Published by Pearson. 2018.–1168 pages.
- 4. Прэтт У. Цифровая обработка изображений: В 2-х книгах. Перевод с англ. М.: Мир, 1983. 790 с.
  - 5. Смирнова Н.В., Пономарева О.В. Векторная и

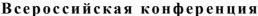
- спектральная обработка сигналов в музыкальной акустике методом параметрического дискретного преобразования Фурье // Цифровая обработка сигналов, 2019. № 2. С. 3-11.
- 6. Пономарева О.В., Пономарев А.В. Быстрый метод горизонтальной скользящей пространственно-частотной обработки // Интеллектуальные системы в производстве. 2019. Т. 17.-№ 2. С. 81-87.
- 7. Пономарев А.В. Основы теории двумерной цифровой обработки сигналов в базисах Фурье с варьируемыми параметрами // Цифровая обработка сигналов, 2019. № 2. С. 12-20.
- 8. Пономарев В.А, Пономарева О.В. Инвариантность текущего энергетического Фурье-спектра действительных дискретных сигналов на конечных интервалах // Технология и конструирование в электронной аппаратуре. 2014. № 1. С. 15-22.
- 9. Пономарев В.А, Пономарева О.В., Пономарева Н.В. Метод быстрого вычисления дискретного преобразования Гильберта в частотной области // Современные информационные и электронные технологии, 2014. № 15. С.183-184.
- 10. Пономарева О.В., Пономарев А.В., Пономарева Н.В. Формализованное описание погрешности измерения вероятностных характеристик случайных процессов процессорными измерительными средствами // Современные информационные и электронные технологии, 2013. № 14. С. 90-93.
- 11. Пономарева Н.В., Пономарева О.В., Хворенков В.В. Определение огибающей ангармонического дискретного сигнала на основе преобразования Гильберта в частотной области // Интеллектуальные системы в производстве. 2018. Т. 16. № 1. С. 33-40.
- 12. Пономарев В.А., Пономарева О.В. Тенденции развития дискретных косвенных измерений параметров электрических сигналов // Метрология. 2017. № 1. С. 20-32.
- 13. Пономарева О.В., Пономарев А.В., Пономарева Н.В. Иерархическое морфологическо-информационное описание систем функционального диагностирования объектов // Современные информационные и электронные технологии, 2013. Т. 1. № 14. С. 121-124.
- 14. Пономарева Н.В. Проблемы компьютерной спектральной обработки сигналов в музыкальной акустике // Интеллектуальные системы в производстве. 2018. Т. 16. № 1. С. 26-33.
- 15. Пономарева Н.В. Цифровая спектральная обработка сигналов в музыкальной акустике // DSPA: Вопросы применения цифровой обработки сигналов. 2018. Т. 8. № 2. С. 37-42.
- 16. Пономарев В.А., Пономарева Н.В. Метод и алгоритм выделения музыкально-акустического сигнала из его смеси со случайным дискретным телеграфным сигналом. В сборнике: Перспективные информационные технологии (ПИТ 2018). Труды Международной научнотехнической конференции. Под ред. С.А. Прохорова. 2018. С. 161-164.
- 17. Пономарева Н.В. Предобработка дискретных сигналов при спектральном анализе в системе компьютерной математики MATLAB // Интеллектуальные системы

- в производстве. 2016. № 4(31). С. 32-34.
- 18. Пономарев В.А., Пономарева Н.В. Цифровой спектрально-временной анализ музыкально-акустических сигналов на основе параметрического дискретного преобразования Фурье// В сборнике: Приборостроение в XXI веке 2017. Интеграция науки, образования и производства. Сборник материалов XIII Международной научно-технической конференции. Ижевск, 2018. С. 307-312.
- 19. Пономарева Н.В., Пономарев В.В. Метод быстрого получения прореженных коэффициентов дискретного преобразования Фурье на основе параметрических дискретных экспоненциальных базисов // DSPA: Вопросы применения цифровой обработки сигналов. 2017. Т. 7.- № 1. С. 172-177.
- 20. Пономарева Н.В, Пономарева В.Ю. Локализация спектральных пиков методом параметрического дискретного преобразования Фурье // Интеллектуальные системы в производстве. 2016. № 2(29). С. 15-18.
- 21. Пономарева Н.В, Пономарева В.Ю. Метод измерения частоты сигналов на базе параметрического дискретного преобразования Фурье // DSPA: Вопросы применения цифровой обработки сигналов. 2016. Т. 6. № 2. С. 393-397.
- 22. Батищев В.И., Золин А.Г., Косарев Д.Н., Романеев А.Е. Аппроксимационный подход к решению задач анализа и интерпретации экспериментальных данных // Вестник Самарского государственного университета. Серия: Технические науки. 2006.-№40.-С.57-65.
- 23. Батищев В.И., Мелентьев В.С. Измерительно-моделирующий подход к определению интегральных характеристик периодических сигналов // Известия высших учебных заведений. Электромеханика. 2003. – №6. – С.36-39.
- 24. Батищев В.И., Волков И.И., Золин А.Г. Использование стохастического базиса в задачах восстановления сигналов и изображений // Автометрия.2017. Т. 53.  $\mathbb{N}_2$  4. С. 127-134.
- 25. Батищев В.И., Волков И.И., Золин А.Г. Исследование аппроксимационных свойств функциональных базисов в задачах реконструкции изображений при дистанционном зондировании земли // В сб.: Проблемы управления и моделирования в сложных системах труды XVIII Международной конференции. Под ред.: Е.А. Федосова, Н.А. Кузнецова, В.А. Виттиха. 2016. С. 304-307.
- 26. Prokhorov S.A., Kulikovskikh I.M. Unique Condition for generalized Laguerre Functions to solve pole Position Problem // Signal Processing. 2015. T. 108. C. 25-29.
- 27. Прохоров С.А., Графкин В.В. Структурно-спектральный анализ случайных процессов. Самара, 2010.
- 28. Прозоров Д.Е., Петров Е.П. Быстрый поиск шумоподобных сигналов. Под ред. Е.П. Петрова. Киров, 2006.
- 29. Dudgeon D.E. Multidimensional Digital Signal Processing Prentice Hall, 1995. 406 p.
- 30. Gonzalez R.C., Woods R.E. Digital Image Processing, 4<sup>th</sup> Ed. Published by Pearson. 2018. 1168 p.
- 31. Прэтт У. Цифровая обработка изображений: В 2-х книгах. Перевод с англ. М.: Мир, 1982. 790 с.
- 32. Systems Analysis of Discrete Two-Dimensional Signal Processing in Fourier Basis // Advances in Signal Processing. Theories, Algorithms, and System Control. Editor: Margarita Favorskaya, Lakmi C. Jain. // Springer. 2020.









# «СОВРЕМЕННЫЕ ТЕХНОЛОГИИ ОБРАБОТКИ СИГНАЛОВ» (СТОС-2021)

27 октября – 29 октября 2021 г. Москва

All-Russian conference «MODERN TECHNOLOGIES OF SIGNAL PROCESSING» (STOS-2021)

Уважаемые коллеги!



#### ПРИГЛАШАЕМ ВАС ПРИНЯТЬ УЧАСТИЕ В РАБОТЕ КОНФЕРЕНЦИИ

#### ОРГАНИЗАТОРЫ:

- Институт радиотехники и электроники им. В.А. Котельникова РАН
- Российское научно-техническое общество радиотехники, электроники и связи им. А.С. Попова
- Министерство образования и науки Российской Федерации
- Московский технический университет связи и информатики

#### ПРИ УЧАСТИИ:

- Федеральное агентство по промышленности РФ
- ОАО «Концерн радиостроения «Вега»
- ОАО «Концерн «Созвездие»
- ФГУП «НИИР»
- Владимирский государственный университет
- Московский авиационный институт
- Балтийский федеральный университет им. И. Канта
- Военная академия РВСН им. Петра Великого
- Московский государственный технический университет им. Н.Э. Баумана
- Рязанский государственный радиотехнический университет им. В.Ф. Уткина

- Московский энергетический институт
- Санкт-Петербургский государственный университет телекоммуникаций им. проф. А.М. Бонч-Бруевича
- Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина)
- Санкт-Петербургского государственного электротехнического университета «ЛЭТИ» им. В.И. Ульянова (Ленина)
- Тульский государственный университет
- Ульяновский государственный технический университет
- Ярославский государственный университет

# ПРОГРАММНЫЙ КОМИТЕТ

Председатель: академик РАН Гуляев Юрий Васильевич Заместитель Председателя: академик РАН Фёдоров И.Б. Члены комитета:

академик РАН **Бугаев А.С.,** член-корр. РАН **Зубарев Ю.Б.,** академик РАН **Каляев И.А.** 

#### ОРГАНИЗАЦИОННЫЙ КОМИТЕТ:

Председатель: член-корреспондент РАН Черепенин Владимир Алексеевич Заместитель Председателя: д.т.н., проф. Витязев Владимир Викторович Сопредседатели: д.т.н., проф. Аджемов А.С., д.т.н., проф. Цимбал В.А. Члены Оргкомитета:

доц. Алёшин В.С., проф. Акиншин Н.С., проф. Брюханов Ю.А., проф. Васильев К.К., проф. Калошин В.А., проф. Мамон Ю.И., проф. Митрофанов Д.Г., проф. Пахотин В.А., проф. Постников И.И., доц. Самсонов Г.А., проф. Сперанский В.С., проф. Степанов С.Н., проф. Чиров Д.С., проф. Ямпурин Н.П.

**Организация работы** в форме пленарных и проблемно-тематических заседаний. Пленарные доклады будут представлены по основным направлениям работы Конференции. Доклады, включенные в Программу конференции, будут опубликованы в **Сборниках докладов Конференции** к началу её работы. По опыту прошедших конференций, мероприятие оказалось очень эффективным местом общения для всех участников мероприятия: представителей науки и производства, заказчиков и разработчиков, потребителей и поставщиков, преподавателей и студентов, работодателей и соискателей, научно-технических издательств и читателей.