

ЦИФРОВАЯ ОБРАБОТКА СИГНАЛОВ

3/2015

НАУЧНО - ТЕХНИЧЕСКИЙ ЖУРНАЛ

ЦИФРОВАЯ ОБРАБОТКА ИЗОБРАЖЕНИЙ тематический выпуск:

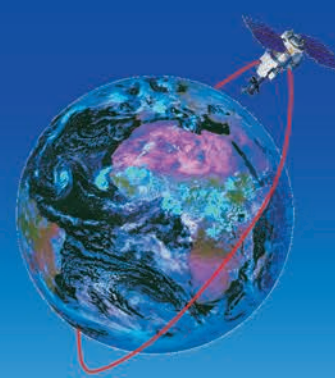
- дистанционное зондирование Земли
- комплексирование снимков Земли
- системы комбинированного видения
- оценка влияния шума на TV изображение
- структурное восстановление изображения
- автоматический контроль геопривязки
- обнаружение игровых эпизодов
- алгоритмы матирования видео
- локализация центров глаз
- сжатие цифровых изображений



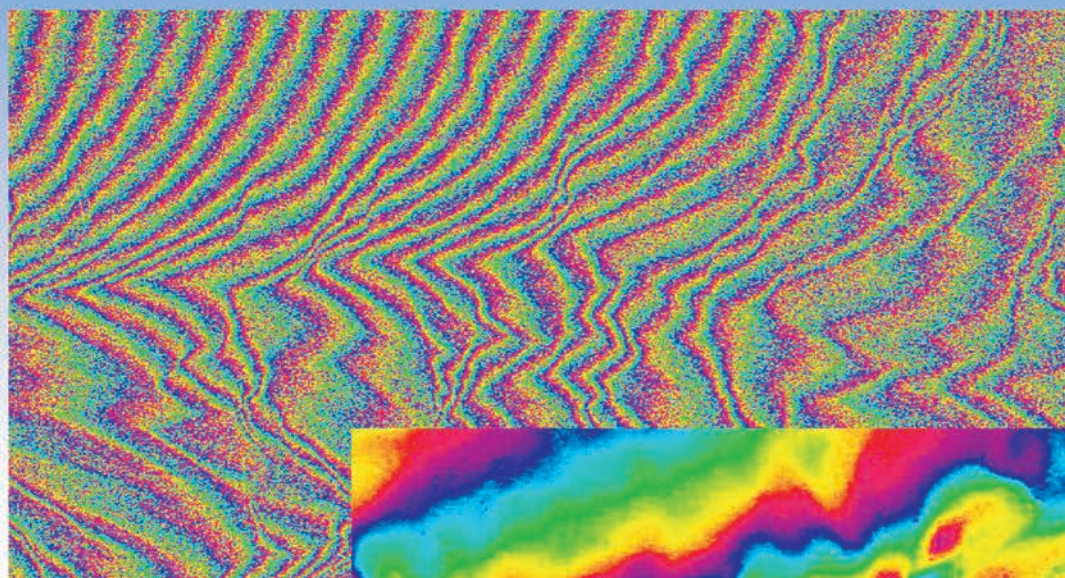


РЯЗАНСКИЙ ГОСУДАРСТВЕННЫЙ РАДИОТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ

НИИ ОБРАБОТКИ АЭРОКОСМИЧЕСКИХ ИЗОБРАЖЕНИЙ

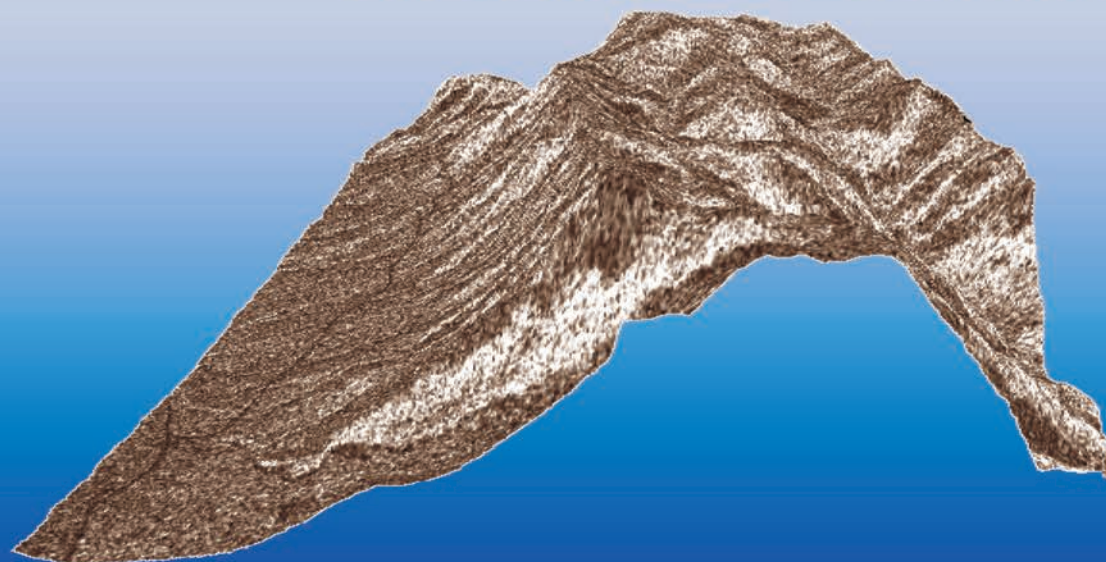
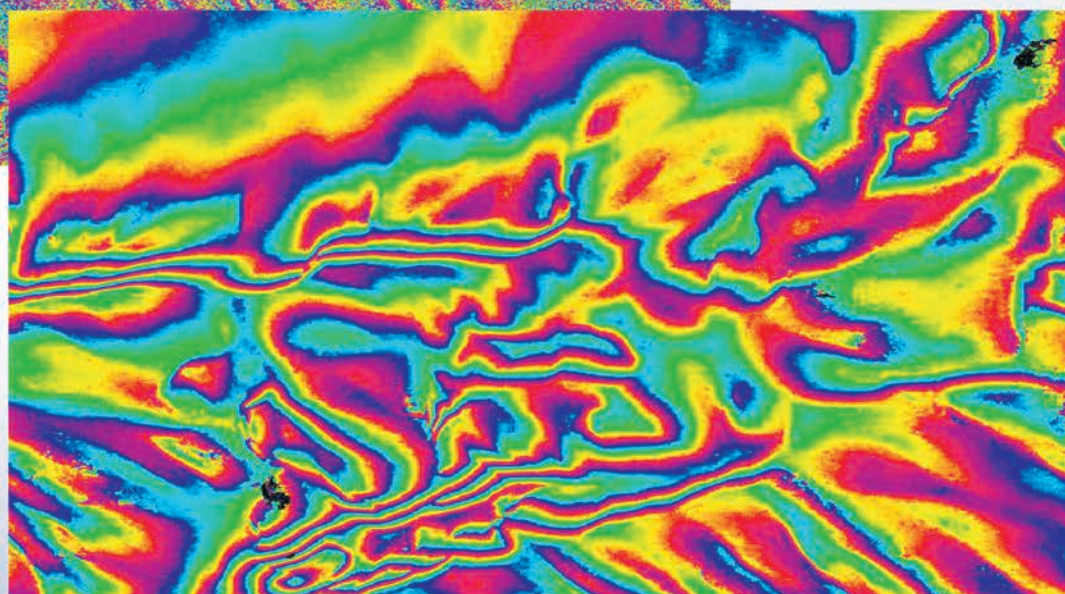


НОВОЕ НАПРАВЛЕНИЕ НАУЧНЫХ ИССЛЕДОВАНИЙ -
ИНТЕРФЕРОМЕТРИЧЕСКАЯ ОБРАБОТКА ИНФОРМАЦИИ ОТ РАДИОЛОКАЦИОННЫХ СИСТЕМ НАБЛЮДЕНИЯ ЗЕМЛИ



*Радиолокационная
интерферограмма
(спутник
«Radarsat-2»)*

*Отфильтрованная
интерферограмма
после вычитания
фазы опорного
рельефа*



*Результат
обработки –
цифровая модель
рельефа
в трехмерном
представлении
с наложенным
радиолокационным
изображением
в качестве
текстуры*

ЦИФРОВАЯ ОБРАБОТКА СИГНАЛОВ

Научно-технический журнал
№ 3/2015

Издается с 1999 года
Выходит четыре раза в год

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

ЗАМЕСТИТЕЛИ ГЛАВНОГО РЕДАКТОРА:

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

РЕДАКЦИОННАЯ КОЛЛЕГИЯ:

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

Адрес редакции:

г. Москва, ул. Авиамоторная, д. 8
Научный центр МТУСИ
Тел.: (8495) 362-42-75
E-mail: rntores@mail.ru
vityazev.v.v@rsreu.ru
http://www.dspsa.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 02.10.15 г.
Формат 60x90/8.

Гарнитура «Arial». Печать офсетная.
Бумага офсетная. Печ.л. 6,5. Тираж 700 экз.

Заказ № 6330. Отпечатано в
ООО НПЦ «Информационные технологии»
Рязань, ул. Островского, д. 21/1
тел.: (4912) 98-69-84

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

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

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

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

В НОМЕРЕ:

- Егошкин Н.А., Еремеев В.В., Макаренков А.А.**
Комплексирование гиперспектральных снимков Земли,
полученных в различных спектральных диапазонах 3
- Алпатов Б.А., Ершов М.Д., Фельдман А.Б.**
Алгоритм обработки изображений
для системы комбинированного видения
летательного аппарата 8
- Балашов О.Е.**
Алгоритм обнаружения наземных препятствий
при маловысотных полетах 15
- Тюкин А.Л., Приоров А.Л.**
Оценка влияния шума на телевизионное изображение
системы позиционирования в помещении 20
- Сильвестрова О.В.**
Оценка эффективности способов цветовой коррекции
для панорамных изображений
с малоразмерными объектами 25
- Кузнецов А.Е., Пресняков О.А., Мятлов Г.Н.**
Структурное восстановление изображений
от многоматричных сканерных датчиков
дистанционного зондирования Земли 29
- Кузнецов А.Е., Пошихонов В.И., Рыжиков А.С.**
Технология автоматического контроля точности геопривязки
спутниковых изображений по опорным
снимкам от КА «Landsat-8» 37
- Петрова К.Ю., Рычагов М.Н., Седунов С.М.**
Обнаружение игровых эпизодов в видеопотоке 43
- Ерофеев М.В., Гитман Ю.А.,
Ватолин Д.С., Федоров А.А.**
Методика объективного сравнения алгоритмов
матирования видео 53
- Хрящев В.В., Приоров А.Л.,
Никитин А.Е., Степанова О.А.**
Итерационный алгоритм локализации центров глаз
на изображении лица 60
- Петров Е.П., Харина Н.Л., Сухих П.Н.**
Метод сжатия цифровых изображений
без спектральных преобразований 65
- Елесина С.И., Ломтева О.А.**
Исследование алгоритмов к
орреляционного совмещения изображений
в системах комбинированного видения 71

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

Digital Signal Processing

Science & Technical Magazine
Issue 3, 2015 year

Is published quarterly since 1999

THE EDITOR-IN-CHIEF:

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

DEPUTY EDITORS-IN-CHIEF:

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

EDITORIAL BOARD:

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

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

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

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

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

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

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

Zubarev R.V.,

Kirpichnikov A.P.,

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

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

Mistyukov V.G.,

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

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

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

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

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

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

Editorial office address:

Aviamotornaya, 8, Moscow, Russia

Research Center of MTUCI

Phone: (8495) 362-42-75

E-mail: rntores@mail.ru,

vityazev.v.v@rsreu.ru

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

Publisher:

Russian A.S. Popov Society for Radioengineering, Electron-
ics & Communications

Computer makeup:

I.A. Blagodarova

FOUNDERS:

InSys

Radio Research and Development Institute

CONTENTS:

Egoshkin N.A., Ereemeev V.V., Makarenkov A.A.

*Fusion of hyperspectral images of the Earth,
acquired in different spectral ranges*3

Alpatov B.A., Ershov M.D., Feldman A.B.

*Image processing algorithm
for combined vision system of aircraft*8

Balashov O.E.

Obstacle detection algorithm at low-altitude flight.....15

Tyukin A.L., Priorov A.L., Demidov P.G.

*The estimation of the noise influence
on the television image
of the indoor positioning system*20

Silvestrova O.V.

*Color correction methods efficiency evaluation
for panoramic images with small-size objects*.....25

Kuznetsov A.E., Presniakov O.A., Myatov G.N.

*Stitching of remote sensing images
from staggered TDI CCD*29

Kuznetsov A.E., Poshehonov V.I., Ryzgikov A.S.

*Automatic georeference accuracy
control technology based on reference pictures
from the observation satellite «Landsat-8»*37

Petrova X.Y., Rychagov M.N., Sedunov S.M.

Field game episode detection in video sequences43

Erofeev M., Gitman Y., Vatolin D., Fedorov A.

*Objective video matting methods'
comparison technique*53

Khryashchev V., Priorov A.,

Nikitin A., Stepanova O.

*Iterative algorithm for the eye centers
localization on facial image*60

Petrov E.P., Kharina N.L., Sukhikh P.N.

*Method of digital images compression without
spectral transformations*65

Elesina S.I., Lomteva O.A.

*Research of the correlation
images combining algorithms
in combined vision systems*71

Subscription index: 82185 ("Rospechat")

ISSN: 1684-2634 9 771684 263005

УДК 004.932

КОМПЛЕКСИРОВАНИЕ ГИПЕРСПЕКТРАЛЬНЫХ СНИМКОВ ЗЕМЛИ, ПОЛУЧЕННЫХ В РАЗЛИЧНЫХ СПЕКТРАЛЬНЫХ ДИАПАЗОНАХ

Егошкин Н.А., к.т.н., в.н.с. НИИ «Фотон» Рязанского государственного радиотехнического университета (РГРТУ), e-mail: foton@rsreu.ru;

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

Макаренков А.А., аспирант НИИ «Фотон» РГРТУ.

FUSION OF HYPERSPECTRAL IMAGES OF THE EARTH, ACQUIRED IN DIFFERENT SPECTRAL RANGES

Egoshkin N.A., Ereemeev V.V., Makarenkov A.A.

Creation of the unified hypercube via fusion of the imagery from separate hyperspectral devices is discussed. Fusion is performed by linear filtering with filter estimation based on the processing images.

Keywords: hyperspectral images, fusion, spatial resolution increasing, image filtering

Ключевые слова: гиперспектральные изображения, комплексирование, повышение пространственного разрешения, фильтрация изображений.

Введение

Гиперспектральная съемка – это новое перспективное направление дистанционного зондирования Земли. В ее основе лежит принцип расщепления отраженной от поверхности Земли лучистой энергии на десятки и сотни потоков, каждый из которых соответствует очень узкому спектральному диапазону, а в целом, регистрируется серия изображений в узких соприкасающихся диапазонах, так называемый гиперкуб, [1, 2, 3]. В результате расщепления регистрируемого излучения энергия полезного видеосигнала многократно уменьшается. В тоже время актуальным вопросом является повышение спектрального разрешения гиперспектрометра путем увеличения числа спектральных каналов. В разработках гиперспектральной съемочной аппаратуры прослеживается тенденция охватить в одном приборе съемку в спектральном диапазоне 0,4 – 2,5 мкм (видимый, ближний ИК и ИК-диапазоны) [4]. Поэтому в практике построения гиперспектрометров получило развитие решение, когда достаточно широкая спектральная область съемки разбивается на ряд частично перекрывающихся диапазонов. Излучение в каждом из этих диапазонов независимо регистрируется с помощью отдельного оптико-электронного тракта, состоящего из оптического узла, диспергирующего элемента и оптико-электронного преобразователя (ОЭП) [4]. При этом все ОЭП выполняют съемку одного и того же участка земной поверхности, но в различных спектральных диапазонах. По такому принципу, например, построены гиперспектрометры, установленные на КА «Ресурс-П» [5] и КА EarthObserver-1 [6]. В результате съемки подобной аппаратурой формируется несколько гиперкубов (каждый в определенном спектральном диапазоне), которые отличаются радиометрическими и геометрическими параметрами,

Рассматриваются вопросы формирования единого гиперспектрального снимка Земли путем объединения информации, полученной от различных видеотрактов гиперспектральной аппаратуры. Объединение производится на основе линейной фильтрации с оценкой параметров фильтра непосредственно по видеоданным.

рами, а также пространственным разрешением.

Стоит задача объединения этих данных в единый гиперкуб, охватывающий все спектральные диапазоны исходных гиперкубов. В качестве примера на рис. 1 слева приведены два гиперкуба до радиометрического и геометрического совмещения, а справа – результат совмещения. Задача геометрического и радиометрического совмещения исходных гиперкубов достаточно освещена в литературе, например в работах [3, 7, 8]. Поэтому в настоящей статье эти вопросы не рассматриваются.

Мало исследованным вопросом является приведение гиперкубов, полученных от различных ОЭП, к единому пространственному разрешению. При этом ставится задача получения объединенного гиперкуба с максимально возможным разрешением. Предпосылкой к такой постановке вопроса является тот факт, что в наборе исходных гиперкубов есть гиперкуб с наилучшим разрешением и он может быть использован для нахождения параметров корректирующего фильтра. Решение этой задачи представлено в настоящей статье. Заметим, что только после выполнения операций геометрического и радиометрического совмещения и выравнивания пространственного разрешения исходных гиперкубов становится возможным выполнение различных видов обработки данных объединенного гиперкуба [9].

Не снижая общности, далее будем рассматривать случай, когда гиперспектрометр включает два ОЭП. Полученные результаты легко распространяются на случай гиперспектральной съемки более чем двумя ОЭП. Предложенная ниже технология решения поставленной задачи является развитием результатов, полученных в работе [10].

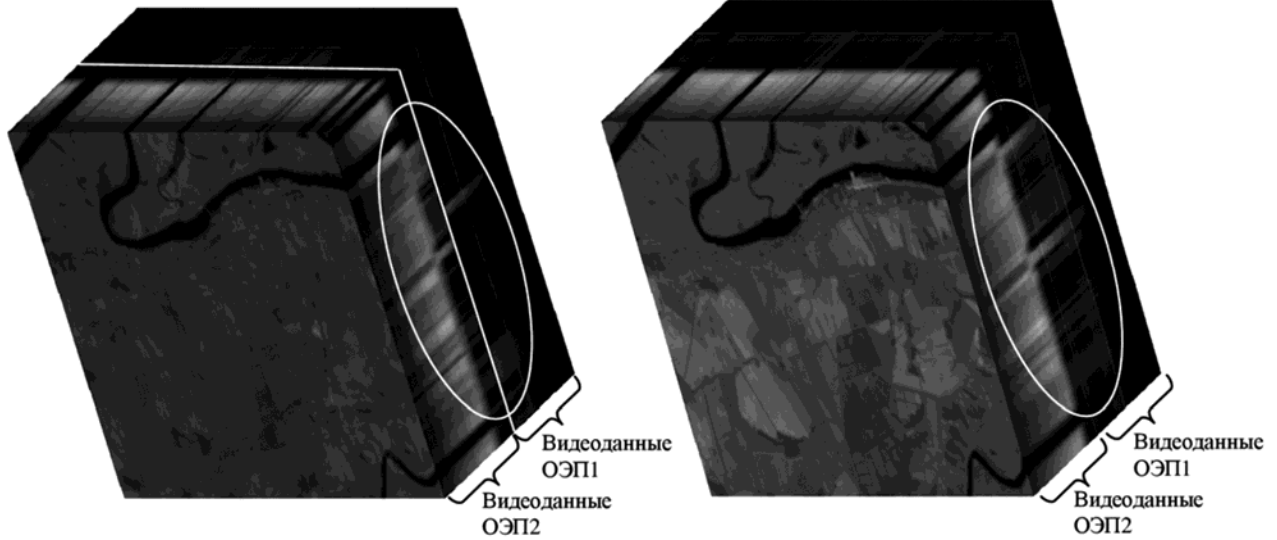


Рис. 1. Формирование единого гиперкуба данных

Формальная постановка задачи

Будем считать, что выполнено радиометрическое и геометрическое совмещение гиперкубов. Рассмотрим два гиперспектральных изображения $A(x, y, \lambda)$, $\lambda \in \Lambda_A$ и $B(x, y, \lambda)$, $\lambda \in \Lambda_B$, сформированных двумя ОЭП со спектральными диапазонами Λ_A и Λ_B . Выберем из этих гиперкубов два изображения $A(x, y)$ и $B(x, y)$, наиболее близкие по длине волны λ (в результате чего переменную λ можно опустить):

$$\begin{aligned} A(x, y) &= S(x, y) \otimes H_A + N_A(x, y), \\ B(x, y) &= S(x, y) \otimes H_B + N_B(x, y), \end{aligned} \quad (1)$$

где $S(x, y)$ – истинное изображение подстилающей поверхности; H_A , H_B – функции рассеяния точки видеотрактов; $N_A(x, y)$, $N_B(x, y)$ – независимые случайные процессы (шум) с нулевым средним; \otimes – операция свертки.

Ставится задача коррекции изображений $A(x, y)$ и $B(x, y)$ так, чтобы $A(x, y) \approx B(x, y)$. Например, можно поставить задачу нахождения линейного фильтра F , обеспечивающего для любого S равенство:

$$S(x, y) \otimes H_A \otimes F = S(x, y) \otimes H_B, \quad (2)$$

т.е. изображение A корректируется с использованием F так, чтобы при отсутствии шумов оно совпало с изображением B .

Задача коррекции различий видеотрактов легко решается, если известны H_A и H_B . Однако нахождение H_A и H_B с достаточной точностью является гораздо более трудной задачей, чем нахождение F , поскольку $S(x, y)$ не известно. Поэтому актуально определение F , не требующее определения H_A и H_B . С учетом (1) условие (2) можно записать как

$$A(x, y) \otimes F - B(x, y) = N_A(x, y) \otimes F - N_B(x, y). \quad (3)$$

Поиск фильтра F можно осуществить в простран-

ственной или спектральной областях. Далее приводятся подходы по определению F в обеих постановках.

Спектральный подход

В спектральной области выражение (3) принимает вид

$$\begin{aligned} \hat{A}(\omega, \xi) \hat{F}(\omega, \xi) - \hat{B}(\omega, \xi) &= \\ = \hat{N}_A(\omega, \xi) \hat{F}(\omega, \xi) - \hat{N}_B(\omega, \xi), \end{aligned} \quad (4)$$

где $\hat{A}(\omega, \xi)$, $\hat{F}(\omega, \xi)$, $\hat{B}(\omega, \xi)$, $\hat{N}_A(\omega, \xi)$, $\hat{N}_B(\omega, \xi)$ – спектры Фурье соответствующих изображений и сигналов. Поэтому

$$\hat{F}(\omega, \xi) = \frac{\hat{B}(\omega, \xi) - \hat{N}_B(\omega, \xi)}{\hat{A}(\omega, \xi) - \hat{N}_A(\omega, \xi)}. \quad (5)$$

Реализация шумовых процессов в (5) не известна. На практике можно оценить $\hat{F}(\omega, \xi)$ как

$$\hat{F}^*(\omega, \xi) = \hat{B}(\omega, \xi) / \hat{A}(\omega, \xi), \quad (6)$$

однако при этом из-за действия шумов возникает ошибка.

Определим относительную погрешность числителя и знаменателя в (6) как

$$\begin{aligned} \varepsilon_A(\omega, \xi) &= \sigma [N_A(\omega, \xi)] / |S(\omega, \xi) H_A(\omega, \xi)|, \\ \varepsilon_B(\omega, \xi) &= \sigma [N_B(\omega, \xi)] / |S(\omega, \xi) H_B(\omega, \xi)|, \end{aligned} \quad (7)$$

где $\sigma[\cdot]$ – СКО, а $|S(\omega, \xi) H_A(\omega, \xi)|$ и $|S(\omega, \xi) H_B(\omega, \xi)|$ – амплитуды «полезных» сигналов $S \otimes H_A$ и $S \otimes H_B$ на частоте (ω, ξ) . Для практического использования (6) необходимо, чтобы $\varepsilon_A(\omega, \xi) \ll 1$ и $\varepsilon_B(\omega, \xi) \ll 1$. При этом вместо $\varepsilon_A, \varepsilon_B$ допустимо использовать их оценки:

$$\begin{aligned} \varepsilon_A^*(\omega, \xi) &= \sigma^* [N_A(\omega, \xi)] / |A(\omega, \xi)|, \\ \varepsilon_B^*(\omega, \xi) &= \sigma^* [N_B(\omega, \xi)] / |B(\omega, \xi)|, \end{aligned} \quad (8)$$

где оценки СКО шума $\sigma^* [N_A(\omega, \xi)]$ и $\sigma^* [N_B(\omega, \xi)]$ можно найти непосредственно по изображениям A, B как это, например, выполнено в работе [11].

Согласно (7) для достоверной оценки (6) необходимо, чтобы:

а) спектральная плотность изображения подстилающей поверхности $|S(\omega, \xi)|$ на частоте (ω, ξ) была достаточно высокой;

б) модули коэффициентов пропускания $H_A(\omega, \xi)$ и $H_B(\omega, \xi)$ видеотрактов ОЭПов, формирующих изображения А и В на частоте (ω, ξ) , были достаточно высокими.

В частности, если $S(\omega, \xi) = 0$, или $H_A(\omega, \xi) = 0$, или $H_B(\omega, \xi) = 0$, то достоверное определение $\hat{F}(\omega, \xi)$ невозможно.

Для реальных изображений подстилающей поверхности $S(\omega, \xi)$ может иметь сложный и достаточно непредсказуемый вид. Например, изображение городской плановой застройки может иметь хорошо выраженный период, в результате $|S(\omega, \xi)|$ окажется недопустимо малым для всех частот, кроме гармоник основного периода, и как следствие, оценка (6) для большинства частот будет недостоверной. Качественную оценку (6) можно получить только путем съемки специальных высококонтрастных калибровочных изображений с широким, достаточно равномерным спектром, или путем подбора изображений подстилающей поверхности с аналогичными свойствами, исходя из условий $\varepsilon_A^*(\omega, \xi) \ll 1$ и $\varepsilon_B^*(\omega, \xi) \ll 1$ для всех (ω, ξ) .

Таким образом, оценку F можно получить, если доступны несколько реализаций изображений A_i, B_i , соответствующих различным сюжетам S_i . Равенство (4) выполняется для каждой реализации, а математическое ожидание правой части (4) равно нулю, поэтому оптимальная оценка $\hat{F}^*(\omega, \xi)$ определяется из условия

$$\sum_i \left| \hat{A}_i(\omega, \xi) \hat{F}^*(\omega, \xi) - \hat{B}_i(\omega, \xi) \right|^2 = \min. \quad (9)$$

По МНК находим решение (9):

$$\hat{F}^*(\omega, \xi) = \frac{\sum_i \hat{B}_i(\omega, \xi) \overline{\hat{A}_i(\omega, \xi)}}{\sum_i \left| \hat{A}_i(\omega, \xi) \right|^2}, \quad (10)$$

где $\overline{\hat{A}_i(\omega, \xi)}$ – комплексно сопряженное для числа $\hat{A}_i(\omega, \xi)$. Используя (10), можно обобщить выражение (8) для нахождения достоверности определения $\hat{F}^*(\omega, \xi)$:

$$\varepsilon_A^*(\omega, \xi) = \sigma^* [N_A(\omega, \xi)] / \sqrt{\sum_i |A(\omega, \xi)|^2}, \quad (11)$$

$$\varepsilon_B^*(\omega, \xi) = \sigma^* [N_B(\omega, \xi)] / \sqrt{\sum_i |B(\omega, \xi)|^2}.$$

Пространственный подход

В этом случае задачу поиска F можно сформулировать как

$$A(x, y) \otimes F^* - B(x, y) \rightarrow \min. \quad (12)$$

При этом полагается, что F^* – функция с финитным носителем, которая может быть представлена в виде «окна» размером $N \times N$. Из выражения (12) сформируем систему линейных уравнений для фрагмента $m \times n$ на изображениях А и В:

$$\begin{cases} A(x_1, y_1) \otimes F^* = B(x_1, y_1), \\ \dots \\ A(x_m, y_1) \otimes F^* = B(x_m, y_1), \\ \dots \\ A(x_m, y_n) \otimes F^* = B(x_m, y_n). \end{cases} \quad (13)$$

Решая систему линейных уравнений (13) найдем параметры оконного фильтра F^* , удовлетворяющего условию (12). Для адекватной оценки F^* при наличии шумов необходимо чтобы $m \cdot n \gg N \cdot N$.

Недостаток подобного подхода состоит в том, что он практически применим только для малых N . В тоже время согласно (2) носитель $F^*(x, y)$ существенно зависит от функций рассеяния точки видеотрактов H_A, H_B . И если функция H_A такова, что решение задачи совмещения видеотрактов включает задачу инверсной фильтрации, носитель $F^*(x, y)$ может быть значительным, даже если H_A, H_B имеют достаточно малый носитель. В этом случае более эффективно решение задачи в однородной постановке: будем искать две функции $F^*(x, y), G^*(x, y)$ с достаточно малым носителем из условия

$$\sum_{(x, y)} \left(A(x, y) \otimes F^*(x, y) - B(x, y) \otimes G^*(x, y) \right)^2 = \min. \quad (14)$$

При этом на функции $F^*(x, y), G^*(x, y)$ необходимо наложить ограничения, чтобы избежать тривиальных решений. В качестве ограничения можно использовать условие:

$$\left| \hat{F}^*(\omega, \xi) \right|^2 + \left| \hat{G}^*(\omega, \xi) \right|^2 = 1, \quad (15)$$

где $\omega = \pi i / N, \xi = \pi j / N, i = \overline{1, N}, j = \overline{1, N}$.

Такое ограничение гарантирует, что каждая пространственная частота будет учтена при решении (14).

На рис. 2 представлен результат приведения изображения $B(x, y)$, имеющего худшее разрешение, к разрешению изображения $A(x, y)$.

При обработке получен и использован фильтр F^* , имеющий следующий вид:

$$F^* = \begin{pmatrix} -0.156 & 0.149 & -0.014 & -0.081 & 0.023 & 0.079 & -0.118 \\ 0.071 & 0.063 & -0.151 & -0.02 & -0.034 & -0.033 & 0.061 \\ 0.044 & -0.167 & 0.043 & 0.451 & -0.184 & 0.044 & -0.015 \\ -9.8 \times 10^{-3} & -0.323 & 0.378 & 0.834 & -0.159 & 0.181 & -0.089 \\ -6.9 \times 10^{-3} & -0.246 & 0.194 & 0.509 & -0.177 & 0.051 & -0.076 \\ 0.159 & 8.1 \times 10^{-3} & -0.136 & -8.9 \times 10^{-3} & -0.123 & 5.8 \times 10^{-3} & 0.102 \\ -0.202 & 0.186 & -0.086 & -0.062 & 0.044 & 0.055 & -0.101 \end{pmatrix}$$

Графически фильтр F^* представлен на рис. 3.

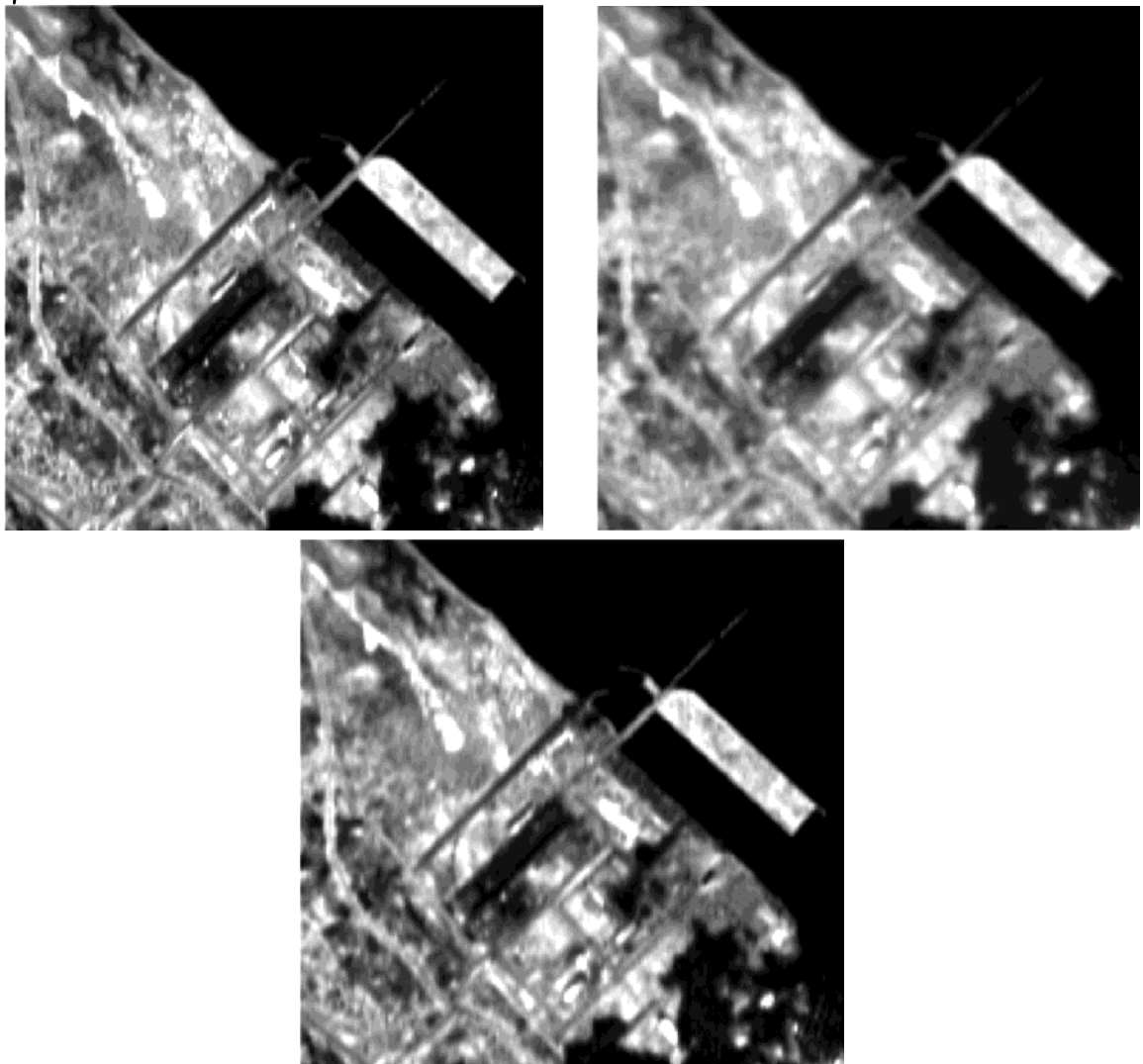


Рис. 2: сверху – изображение $A(x, y)$; в центре – изображение $B(x, y)$, имеющее худшее разрешение; внизу – результат обработки $B(x, y)$ фильтром F^*

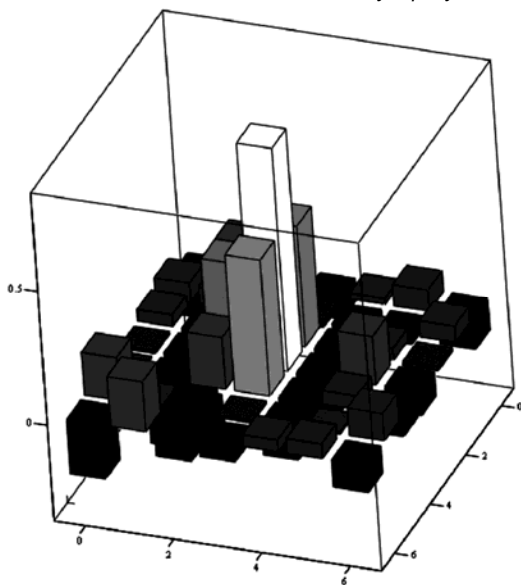


Рис. 3. Графическое представление фильтра F^*

Заключение

Для серии изображений $A(x, y)$ и $B(x, y)$ (на рис. 2 сверху и в центре представлена одна из реализаций) с

использованием подхода, изложенного в работе [11], выполнена достоверная оценка пространственных разрешений исходных изображений и результата их обработки. На достаточно представительном статистическом материале от гиперспектрометра спутника «Ресурс-П» установлено:

1) функции рассеяния точки реальных видеотрактов H_A, H_B позволяют оценить восстанавливающий фильтр F^* с достаточной точностью при размерах «окна» 7×7 ;

2) при относительном различии пространственных разрешений исходных изображений 35% после обработки этот показатель сократился до 14%.

3) фильтрация с помощью F^* дополнительно позволила существенно снизить геометрические рассогласования между исходными изображениями $A(x, y)$ и $B(x, y)$.

Литература

1. Шовенгердт Р.А. Дистанционное зондирование. Модели и методы обработки изображений. М.: Техносфера, 2012. 560 с.
2. Еремеев В.В., Макаренков А.А., Москвитин А.Э., Мятлов Г.Н. Повышение информативности материалов

съемки Земли за счет объединения гиперспектральной информации с данными от других съемочных систем // Цифровая обработка сигналов. 2013. №4. С. 37–41.

3. Современные технологии обработки данных дистанционного зондирования Земли / Под ред. В.В. Еремеева. – М.: ФИЗМАТЛИТ, 2015. – 460 с.

4. Архипов С.А., Кравченко С.О., Ли А.В., Линько В.М., Морозов С.А. Схемные решения оптического тракта изображающих спектрометров // Матер. научн. техн. конф. «Гиперспектральные приборы и технологии». М: ОАО «Красногорский завод им. С.А. Зверева», 2013. С. 44–45.

5. Архипов С.А., Ляхов А.Ю., Тарасов А.П. Работы ОАО «Красногорский завод им. С.А. Зверева» по созданию гиперспектральных приборов дистанционного зондирования // Матер. научн. техн. конф. «Гиперспектральные приборы и технологии». М: ОАО «Красногорский завод им. С.А. Зверева», 2013. С. 25–30.

6. Mark A. Folkman, Jay Pearlman, Lushalan B. Liao, Peter J. Jarecke, «EO-1/Hyperion hyperspectral imager design, development, characterization, and calibration», Proc. SPIE 4151, Hyperspectral Remote Sensing of the Land and Atmosphere, 40 (February 8, 2001).

7. Кузнецов А.Е., Светелкин П.Н. Формирование цветных снимков по спутниковым изображениям среднего и высокого пространственного разрешения // Цифровая обработка сигналов. 2009. №3. С. 36–40.

8. Архипов С.А., Бутякин Н.А., Ли А.В., Линько В.М., Лузганова М.Ю. Предварительная обработка гиперспектральной видеоинформации ГСА РП // Матер. научн. конф. «Гиперспектральные приборы и технологии». М: ОАО «Красногорский завод им. С.А. Зверева», 2013. С. 132.

9. Ахметов Р.Н., Везенов В.И., Еремеев В.В., Стратилатов Н.Р., Юдаков А.А. Модели формирования и некоторые алгоритмы обработки гиперспектральных изображений // Исследование Земли из космоса. 2014. №1. С. 17–28.

10. Егошкин Н.А., Еремеев В.В. Коррекция смаза изображений в системах космического наблюдения земли // Цифровая обработка сигналов. №4. 2010. С. 28 – 32.

11. Еремеев В.В., Князьков П.А., Козлов Е.П., Москвитин А.Э. Технологии анализа и оценки качества космических изображений // Цифровая обработка сигналов. 2010. №3. С. 30–37.

НОВЫЕ КНИГИ

Горбунов Ю.Н., Лобанов Б.С., Куликов Г.В.

Введение в стохастическую радиолокацию: учеб. пособие. – М.: Горячая линия – Телеком, 2015. – 376 с.

Рассмотрены основы радиолокации, вопросы реализации когерентно-импульсных систем радиолокации и их компонент – цифровых систем обнаружения и измерения параметров радиолокационных сигналов систем селекции движущихся целей, обеспечивающих выделение полезных сигналов от целей на фоне коррелированных помех по доплеровским и угловым признакам (пространственным частотам). В отличие от традиционного изложения материала, упор сделан на стохастические методы обработки сигналов и информации.

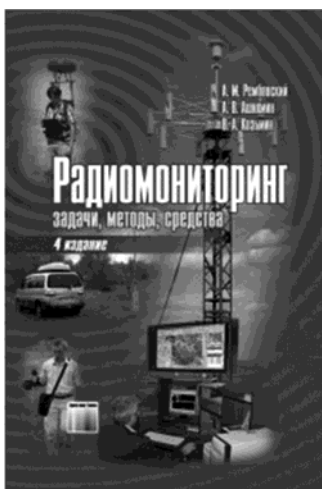
Для студентов, обучающихся по специальности 210601 – «Радиоэлектронные системы и комплексы», а также других радиотехнических и инфокоммуникационных специальностей.

Рембовский А.М., Ашихмин А.В., Козьмин В.А.

Радиомониторинг: задачи, методы, средства: научн. издание под ред. Рембовского А.М. – М.: Горячая линия – Телеком, 2015. – 640 с., 4-е изд., с испр.

Рассмотрены типовые задачи, решаемые с помощью средств радиомониторинга в самых различных областях – от управления использованием радиочастотного спектра до контроля радиообстановки при проведении контртеррористических операций. Изложены теоретические и практические вопросы построения и функционирования радиотехнических систем радиомониторинга и технической защиты информации, в том числе автоматизированных стационарных и мобильных систем пеленгования, определения местоположения источников радиоизлучения, измерения параметров сигналов, видов радиоизлучений и утечек информации. Рассмотрены математические методы и программное обеспечение для выявления и анализа радиоизлучений, измерения параметров сигналов и помех, определения положения источников радиосигналов и помех. Представлен анализ особенностей построения современных одноканальных и многоканальных широкополосных цифровых панорамных измерительных радиоприемников и радиопеленгаторов, в том числе и CDMA.

Для специалистов в области радиомониторинга, операторов и руководителей радиоконтрольных служб, а также студентов и аспирантов.



АЛГОРИТМ ОБРАБОТКИ ИЗОБРАЖЕНИЙ ДЛЯ СИСТЕМЫ КОМБИНИРОВАННОГО ВИДЕНИЯ ЛЕТАТЕЛЬНОГО АППАРАТА

Алпатов Б.А., д.т.н., профессор, заведующий кафедрой автоматики и информационных технологий в управлении Рязанского государственного радиотехнического университета (РГРТУ), e-mail: aitu@rsreu.ru;

Ершов М.Д., инженер кафедры автоматики и информационных технологий в управлении РГРТУ, e-mail: aitu@rsreu.ru;

Фельдман А.Б., к.т.н., доцент кафедры автоматики и информационных технологий в управлении РГРТУ, e-mail: aitu@rsreu.ru.

IMAGE PROCESSING ALGORITHM FOR COMBINED VISION SYSTEM OF AIRCRAFT

Alpatov B.A., Ershov M.D., Feldman A.B.

Actual problems arising in the development of combined vision systems for aircraft are considered. The algorithm for heterogeneous image registration is offered. This algorithm includes edge detection and estimation of the parameters of geometric transformations of real and synthesized images. Results of experimental research of proposed algorithm are shown.

Keywords: combined vision system, image registration, edge detection, fuzzy clusterization, geometric transformations, Fourier transform.

Ключевые слова: система комбинированного видения, совмещение изображений, выделение границ, нечеткая кластеризация, геометрические преобразования, преобразование Фурье.

Введение

Важную роль в повышении безопасности полетов играет использование специальных технических устройств, способных предупредить экипаж летательного аппарата о возможном столкновении, указать ему местоположение важных ориентиров: взлетно-посадочной полосы, рек и других водоемов, сети автомобильных и железных дорог. К устройствам такого рода относят, в частности, бортовые системы комбинированного видения, осуществляющие совмещение телевизионных изображений местности с данными, получаемыми на основе цифровой топографической карты [1]. Общий принцип работы такой системы, рассматриваемый в рамках данной статьи, проиллюстрирован на рис. 1.

Рассмотрены актуальные проблемы, возникающие при разработке систем комбинированного видения летательных аппаратов. Предложен алгоритм совмещения разнородных изображений. Данный алгоритм включает выделение границ на натурном изображении и оценивание параметров геометрических преобразований натурного и синтезированного изображений. Представлены результаты экспериментальных исследований предложенного алгоритма.

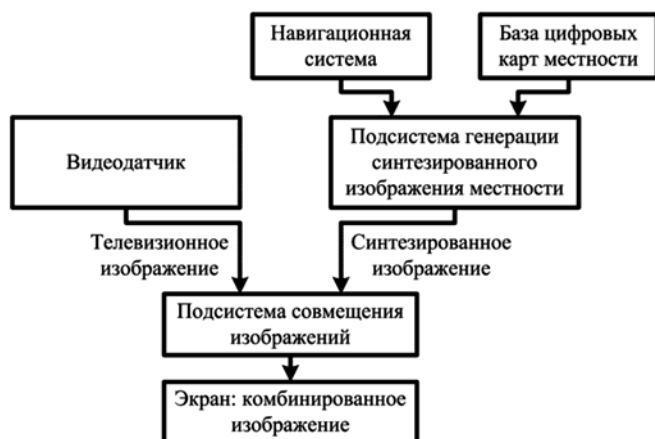
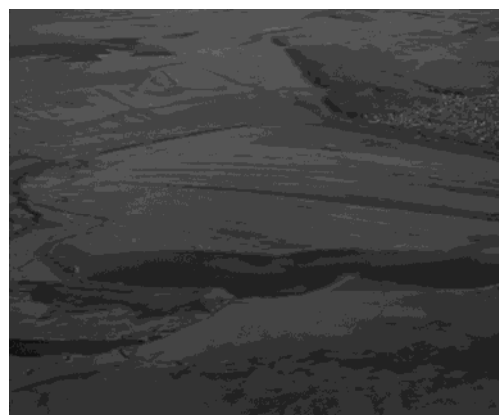
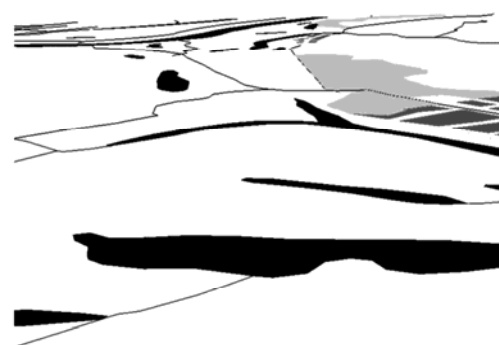


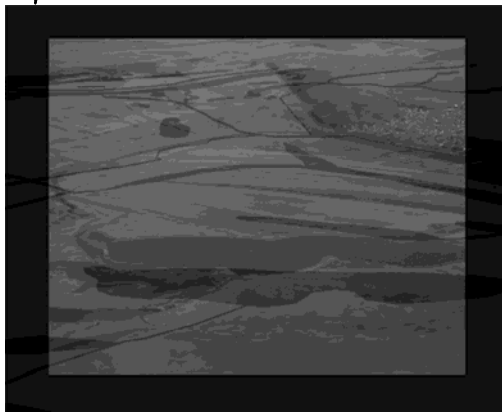
Рис. 1. Структура системы комбинированного видения



а)



б)



в)

Рис. 2. Примеры изображений местности: а – телевизионное; б – на основе цифровой карты местности; в – комбинированное

Вместе с изображением, полученным от видеодатчика, используется синтезированное изображение, построенное на основе цифровой карты местности в соответствии с текущими навигационными параметрами летательного аппарата. Примеры телевизионного и синтезированного изображений приведены на рис. 2. Однако при создании комбинированного изображения зачастую возникает рассогласование натурального и синтезированного изображений (рис. 2, в), связанное с наличием погрешностей измерения положения летательного аппарата и углов его ориентации [2]. Несмотря на существенный прогресс в развитии технических средств воздушной навигации, последние не всегда могут обеспечить достаточную точность и надежность измерения навигационных параметров летательного аппарата.

Таким образом, при разработке системы комбинированного видения необходимо решить задачу геометрического совмещения натурального и синтезированного изображений.

Выполнив точную привязку текущего изображения к цифровой карте местности, можно распознать наземные объекты, обозначенные на карте, и отобразить сведения о них на комбинированном изображении.

Алгоритм синтеза комбинированного изображения

Для решения проблемы рассогласования изображений в данной статье предлагается алгоритм синтеза комбинированного изображения. Исходя из структуры системы комбинированного видения, для работы предлагаемого алгоритма предоставляются следующие исходные данные:

1. Навигационные параметры летательного аппарата: географические координаты, барометрическая высота, углы ориентации. При этом известны погрешности измерения навигационных параметров.

2. Натурное телевизионное изображение. При этом известны такие характеристики оптического датчика, как разрешение и поле зрения.

3. Цифровая карта местности в векторном формате.

Сопоставляемые изображения имеют разную природу, что в итоге необходимо учитывать. Изображение объектов, сформированное на основе цифровой карты местности, является искусственным, размеры и форма

объектов не меняются в зависимости от погодных, сезонных и других условий, в то время как на естественных изображениях местности, полученных с помощью датчиков системы технического зрения, такие условия играют значительную роль. Таким образом, алгоритм синтеза комбинированного изображения должен решать задачу совмещения двух разнородных изображений, при этом не представляется возможной реализация непосредственного сопоставления натурального и синтезированного изображений путем анализа яркостей точек [3], которое к тому же трудно осуществить в режиме реального времени. Также в данном случае не подходит поиск параметров геометрических преобразований изображений с использованием детекторов особых точек [4], так как данные алгоритмы направлены на решение задач сопоставления однородных изображений.

Для решения задачи оценивания параметров геометрических преобразований изображений предлагается использовать корреляционно-экстремальный подход [5]. Исходные разнородные изображения необходимо преобразовать в пригодную для сопоставления форму, для чего можно выделить особенности или признаки. В качестве особенностей, характерных как для изображения, полученного от датчиков системы технического зрения, так и для изображения на основе цифровой карты местности, выбраны границы топографических объектов: автомобильных и железных дорог, водоемов.

С учетом изложенного, алгоритм синтеза комбинированного изображения содержит следующие шаги:

1. Предварительная обработка натурального изображения и выделение границ характерных линейно-протяженных объектов.

2. Генерация набора контурных изображений на основе цифровой карты местности для различных ракурсов с учетом погрешностей измерения навигационных параметров.

3. Корреляционно-экстремальное сопоставление контурных изображений с целью определения параметров рассогласования изображений.

4. Построение комбинированного изображения с учетом параметров рассогласования натурального и синтезированного изображений.

Далее рассмотрим более подробно шаги предложенного алгоритма.

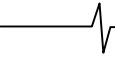
Выделение границ объектов

Проблематике решения задачи выделения границ посвящено большое количество работ [6]. В настоящее время распространено определение операции выделения границ как извлечения точек, в которых имеется значительный перепад яркости относительно соседних точек или областей изображения. В основу разработанного алгоритма выделения границ на изображении, полученном от оптического датчика системы технического зрения, положен метод нечеткой кластеризации [7]. Рассмотрим шаги данного алгоритма:

1. Предобработка телевизионного изображения.

2. Нечеткая кластеризация по методу C -средних с использованием гистограммы яркости изображения.

3. Построение бинарного изображения A на основе



разбиения, полученного на предыдущем шаге.

4. Морфологическое закрытие бинарного изображения A структурным элементом B для устранения неоднородности больших сегментов, выраженной небольшими внутренними «дырками».

5. Оконтуривание полученного на предыдущем шаге изображения и векторизация контуров.

6. Устранение контуров малой протяженности.

Предварительная обработка натурального изображения, как правило, включает подавление шума. Кроме того, в зависимости от параметров датчиков системы технического зрения и условий съемки, с целью улучшения изображения для дальнейшей обработки можно выполнить коррекцию гистограммы яркости.

В простейшем случае для подавления шума используется усредняющий фильтр: каждая точка на выходе является средним арифметическим точек в области заданного размера. Другим линейным фильтром низких частот, который целесообразно применять, является фильтр Гаусса.

Билатеральный фильтр обеспечивает лучшее качество, но требует больших вычислительных затрат. Данный фильтр имеет нелинейный характер и сохраняет границы объектов на изображении. Значение яркости в каждой точке изображения заменяется средневзвешенным значением яркости близлежащих точек. При этом, как правило, используется весовая функция Гаусса, но веса зависят не только от Евклидова расстояния между точками, но также и от других характеристик, например, от яркости точек.

Коррекция гистограммы зачастую применяется для повышения качества изображений. Данная операция позволяет усилить контрастность изображения и повысить детализацию, что в дальнейшем может положительно повлиять на результат выделения границ. В качестве методов коррекции гистограммы яркости можно использовать линейное растяжение гистограммы или выравнивание гистограммы.

На втором шаге алгоритма применяется нечеткая кластеризация по методу C -средних с использованием гистограммы яркости. Кластеризация в широком смысле может быть определена как процесс разделения N элементов данных на заданное число K групп, члены которых подобны друг другу по некоторому признаку. В роли алгоритма кластеризации используется нечеткая кластеризация по методу C -средних (Fuzzy C -Means clusterization) [7]. Данный метод требует задания начальных центров кластеров. Для решения этой задачи предлагается провести начальное определение центров с помощью кластеризации по методу K -средних (K -Means clusterization) [8].

Изначально центры кластеров c_k размещаются на максимально большом расстоянии друг от друга:

$$c_k = b_{\min} + (k-1) \cdot db + \frac{db}{2}, \quad db = \frac{b_{\max} - b_{\min}}{K}, \quad (1)$$

где b_{\min} и b_{\max} – минимальное и максимальное значения яркости изображения, k – номер кластера, $k = \overline{1, K}$, K – число кластеров.

Кластеризация по методу K -средних в рамках разра-

ботанного подхода к выделению границ заключается в проходе по значениям яркостей точек и связывание их с ближайшим центром тяжести с использованием гистограммы. Далее необходимо пересчитывать положение центров тяжести в соответствии с результатом предыдущего шага. Образуется цикл, критерием останова которого является условие $\max \{ |c_k(i) - c_k(i-1)| \} < \varepsilon_k$, то есть оценивается расстояние между центрами кластеров, полученными на предыдущем и текущем шагах. Выполняется поиск минимума целевой функции

$$J = \sum_{n=1}^N \sum_{k=1}^K \|b_{nk} - c_k\|^2, \quad (2)$$

где $\|b_{nk} - c_k\|^2$ – выбранная мера расстояния между элементом данных (яркостью b_{nk}) и центром кластера c_k , к которому данный элемент отнесен, N – разрядность гистограммы.

Расчет центров тяжести кластеров на каждой итерации цикла выполняется в соответствии с выражением

$$c_k = \frac{\sum_{n=1}^N b_{nk} \cdot H(b_{nk})}{\sum_{n=1}^N H(b_{nk})}, \quad (3)$$

где H – гистограмма изображения.

Центры кластеров, определенные с помощью кластеризации по методу K -средних, далее используются в работе алгоритма нечеткой кластеризации по методу C -средних. В ходе решения задачи нечеткой кластеризации осуществляется минимизация целевой функции

$$J_\phi = \sum_{n=1}^N \sum_{k=1}^K u_{nk}^\phi \|b_n - c_k\|^2, \quad 1 \leq \phi < \infty, \quad (4)$$

где ϕ – степень нечеткости, u_{nk} – степень принадлежности элемента данных (яркости b_n) к кластеру k , c_k – центр кластера k .

Нечеткое разбиение осуществляется путем итеративной оптимизации целевой функции (4), при этом производится обновление степеней принадлежности u_{nk} и центров кластеров c_k . Для этого сначала производится расчет расстояний D_{nk} между элементом данных и центрами кластеров:

$$D_{nk} = |b_n - c_k|^{\frac{2}{\phi-1}}. \quad (5)$$

Далее степени принадлежности и центры кластеров рассчитываются в соответствии с выражением

$$u_{nk} = \frac{1}{D_{nk} \cdot \sum_{s=1}^K \frac{1}{D_{ns}}}, \quad c_k = \frac{\sum_{n=1}^N u_{nk}^\phi \cdot H(b_n) \cdot b_n}{\sum_{n=1}^N u_{nk}^\phi \cdot H(b_n)}. \quad (6)$$

Работа цикла прерывается, когда выполняется условие $\max_{nk} \{ |u_{nk}(i) - u_{nk}(i-1)| \} < \varepsilon_c$, где критерий останова ε_c лежит между 0 и 1. С целью уменьшения вычислительных затрат возможно использование условия $\max \{ |c_k(i) - c_k(i-1)| \} < \varepsilon_c$.

Для дальнейшей работы номер кластера k_n , соответствующий значению яркости b_n , принимается равным номеру кластера, при котором достигается $\max\{u_{nk}\}$ – максимум степени принадлежности. В итоге строится таблица поиска номера кластера по значению яркости для последующей разметки кластеров на изображении. На рис. 3 приведен результат кластеризации по гистограмме яркости.

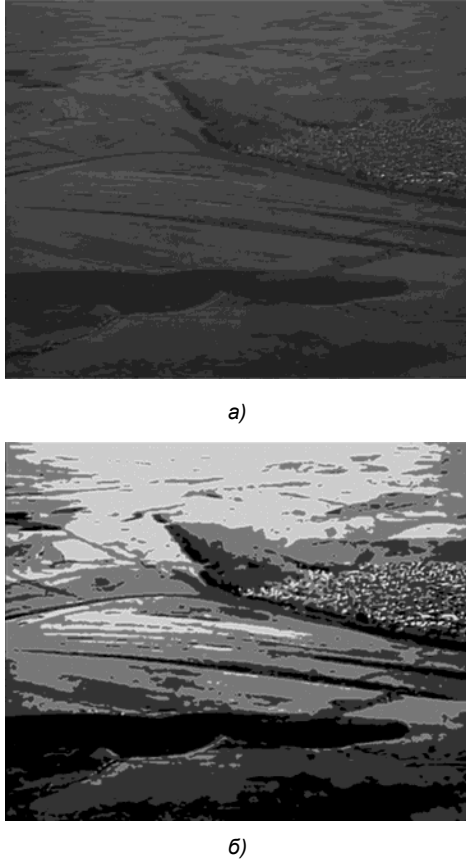


Рис. 3. Исходное изображение (а) и разбиение точек на 4 кластера (б)

Полученное бинарное изображение может содержать незначительные малоразмерные элементы и неоднородности больших сегментов, выраженные небольшими внутренними «дырками», поэтому для улучшения такого изображения предлагается использовать математическую морфологию, а также параметрическую фильтрацию контуров (шаги 4 и 6).

Так как при обработке данных используется глобальная характеристика изображения – его гистограмма яркости, то для повышения точности выделения границ применяется поблочная обработка изображения. Поблочная обработка заключается в выполнении шагов 2, 3 и 4 в каждой области интереса. Размеры и положение областей интереса могут устанавливаться в соответствии с информацией об объектах (таких как дороги, реки и другие водоемы), получаемой из цифровой карты местности. Учитывая использование такой априорной информации, можно полагать, что точки в каждой заданной области интереса относятся к двум кластерам – топографическому объекту и фону. Поэтому для работы алгоритма нечеткой кластеризации число кластеров устанавливается равным двум.

Пример выделения границ на изображении представлен на рис. 4.

Блок-схема алгоритма выделения границ на натурном изображении приведена на рис. 5, где I – исходное натурное изображение, LUT_b – таблица поиска номера кластера по значению яркости, N_p – порог, задающий минимальное число точек контура.

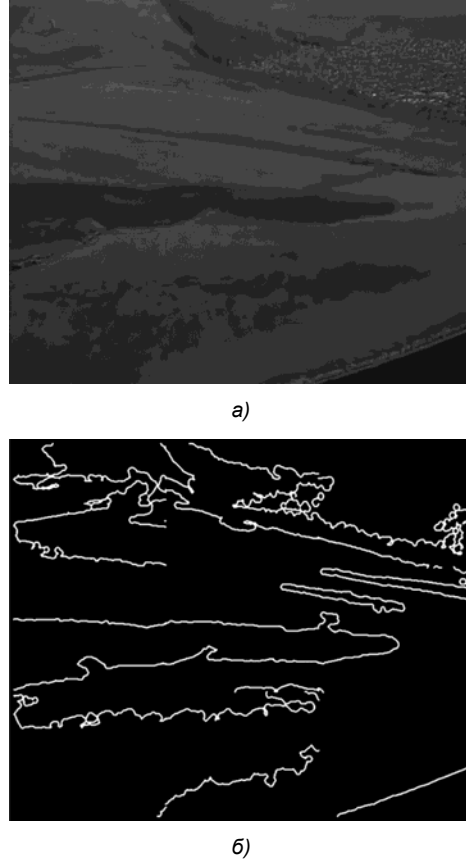


Рис. 4. Исходное изображение (а) и выделенные границы (б)

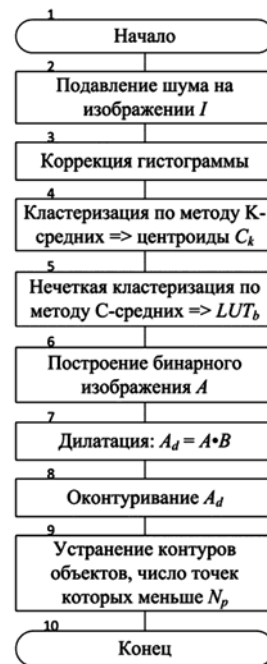


Рис. 5. Блок-схема алгоритма выделения границ на основе нечеткой кластеризации

Оценивание рассогласования натурального и синтезированного изображений

Задачу устранения рассогласования изображений можно рассматривать как процесс отождествления текущего изображения кабины обстановки с одним из множества синтезированных изображений. Построение последних осуществляется с учетом ошибок измерения навигационных параметров летательного аппарата.

Построение контурного изображения местности по цифровой карте производится для каждого ракурса (X, Y, H, Cr, Tn, Kr) , где X и Y – географические координаты, H – барометрическая высота, Cr , Tn и Kr – углы курса тангажа и крена соответственно.

Для каждого из ракурсов необходимо вычислить значение целевой функции, характеризующее степень соответствия изображений. Далее определяется ракурс, при котором было получено максимальное значение данной функции. В качестве целевой функции примем количество ненулевых пикселей бинарного контурного изображения цифровой карты, соответствующих ненулевым пикселям бинарного контурного изображения, полученного в результате выделения границ на натурном изображении.

Построение нового изображения местности и вычисление целевой функции для каждого ракурса при переборе шести навигационных параметров может требовать значительных вычислительных затрат. Более эффективным подходом к решению поставленной задачи является замена перебора углов ориентации Cr , Tn , Kr на оценивание евклидовых преобразований изображений, то есть смещения (α, β) и поворота на угол ϕ . Такая замена допустима в силу того, что при малых углах зрения и малых ошибках измерения углов ориентации летательного аппарата, параметры α , β и ϕ можно считать пропорциональными величинам ошибок измерения углов курса Cr , тангажа Tn и крена Kr .

Исходя из вышесказанного, для решения задачи должны использоваться алгоритмы оценивания параметров геометрических преобразований, позволяющие оценивать смещение (α, β) и поворот на угол ϕ , соответственно при генерации изображений на основе карты местности необходимо расширить поле зрения. Поле зрения расширяется по вертикали и по горизонтали с учетом погрешностей измерения углов тангажа и курса соответственно.

Алгоритм оценки параметров геометрических преобразований использует подход на основе преобразования Фурье, позволяющий снизить вычислительные затраты при реализации корреляционных алгоритмов. Использование данного подхода позволяет определять смещение и поворот изображений на основе свойств преобразования Фурье, отражающих поведение двумерного спектра Фурье [9]. В основе метода оценки параметров смещения и поворота лежат теоремы смещения и вращения для спектров функций двух переменных [5].

Если функция $f_2(x, y)$ получена путём сдвига и поворота функции $f_1(x, y)$, то есть

$$f_2(x, y) = f_1(x \cos \phi + y \sin \phi - \alpha, -x \sin \phi + y \cos \phi - \beta), \quad (7)$$

то их спектры связаны следующим образом:

$$F_2(\omega_x, \omega_y) = E \cdot F_1(\omega_x \cos \phi + \omega_y \sin \phi, -\omega_x \sin \phi + \omega_y \cos \phi),$$

$$E = \exp\{-2\pi j(\omega_x \alpha + \omega_y \beta)\}, \quad (8)$$

где ω_x , ω_y – пространственные частоты, $j = \sqrt{-1}$.

Пусть $M_1(\omega_x, \omega_y)$ и $M_2(\omega_x, \omega_y)$ – амплитудные спектры функций $f_1(x, y)$ и $f_2(x, y)$, то есть

$$M_1(\omega_x, \omega_y) = |F_1(\omega_x, \omega_y)|, \quad M_2(\omega_x, \omega_y) = |F_2(\omega_x, \omega_y)|,$$

тогда из (8) следует

$$M_2(\omega_x, \omega_y) = M_1(\omega_x \cos \phi + \omega_y \sin \phi, -\omega_x \sin \phi + \omega_y \cos \phi). \quad (9)$$

Таким образом, при сдвиге и повороте функции происходит такой же поворот её амплитудного спектра, а сдвиг не оказывает влияния на амплитудный спектр функции. Переходя к амплитудным спектрам, можно освободиться от влияния сдвигов и отслеживать поворот отдельно от других видов преобразований. Для этого удобно в выражении (9) перейти в полярную систему координат: $\omega_x = \rho \cos \theta$, $\omega_y = \rho \sin \theta$. В полярных координатах амплитудные спектры связаны преобразованием сдвига вдоль оси θ :

$$M_2(\rho, \theta) = M_1(\rho, \theta - \phi). \quad (10)$$

Угол ϕ легко находится методом фазовой корреляции [5]. Зная ϕ , можно развернуть функцию $f_1(x, y)$ на угол ϕ и получить две функции, связанные преобразованием сдвига. Применяя вновь метод фазовой корреляции, получим оценки параметров сдвига (α, β) .

Для оптимизации вычислительных затрат необходимо использовать алгоритм быстрого преобразования Фурье (БПФ). Блок-схема алгоритма поиска смещений и поворота с использованием БПФ приведена на рис. 6. На блок-схеме используются следующие обозначения: I – текущее изображение, Et – изображение, относительно которого оцениваются параметры сдвига и поворота. Отметим, что амплитудный спектр дискретного изображения оказывается сильно искаженным вследствие хорошо известного явления Гиббса, которое необходимо компенсировать. После вычисления БПФ исходного изображения выполняется переход к логарифму амплитудного спектра, который в дальнейшем переводится в полярную систему координат. Для определенности будем считать, что перемещение вдоль строки соответствует изменению угла, а перемещение по столбцу – изменению расстояния. Так как необходимо оценить только угол поворота, то с целью сокращения вычислительных затрат вычисление БПФ целесообразно осуществлять только для строк логарифма амплитудного спектра.

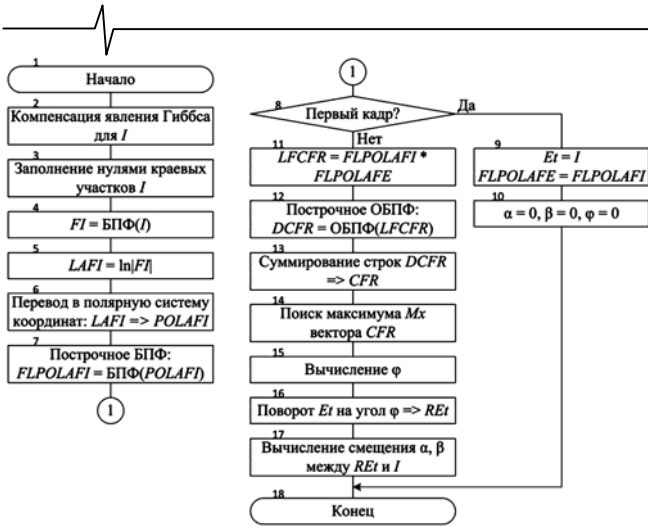


Рис. 6. Блок-схема алгоритма оценивания параметров смещений и поворота на основе БПФ

Экспериментальные исследования

При проведении экспериментальных исследований использовались пять натуральных видеопоследовательностей, полученных при наблюдении с борта летательного аппарата. Для каждого кадра записан набор навигационных параметров. Также использовались цифровые карты местности, соответствующие маршрутам полетов. В ходе исследований оценивался предложенный и реализованный алгоритм обработки изображений для системы комбинированного видения, включающий выделение границ на основе нечеткой кластеризации по методу C-средних и оценивание параметров геометрических преобразований изображений на основе свойств преобразования Фурье. Затем производилось совмещение натурального и синтезированного изображений с учетом найденных параметров рассогласования.

В табл. 1 приведено сравнение среднего времени, затрачиваемого на выделение границ, подготовку изображения местности на основе цифровой карты и оценивание параметров геометрических преобразований при использовании разработанных подходов. Экспериментальные исследования выполнены для следующих интервалов изменения навигационных параметров:

- от -20 м до 20 м с шагом 10 м для координат X и Y;
- от -10 м до 10 м с шагом 5 м для высоты H;
- от -1° до 1° с шагом 1° для углов курса, тангажа и крена.

В данном случае проводилась грубая оценка углов ориентации, что не касается подхода на основе БПФ, так как он обеспечивает максимальную точность, соответствующую размеру элемента разложения телевизионного изображения. Подход на основе полного перебора является неоптимальным по временным затратам по сравнению с другими алгоритмами синтеза комбинированного изображения, поэтому в дальнейшем он не рассматривался.

В табл. 2 приведены результаты для следующих интервалов изменения навигационных параметров:

- от -20 м до 20 м с шагом 10 м для координат X и Y;
- от -10 м до 10 м с шагом 5 м для высоты H;
- от -2° до 2° с шагом 0,125° для углов курса, тангажа и крена.

Использование преобразования Фурье позволяет добиться высокой точности при определении значений углов курса, тангажа и крена. Данный подход стоит применять при больших зонах поиска параметров геометрических преобразований. Однако в случаях, когда высокая точность не требуется, алгоритм на основе преобразования Фурье заметно проигрывает корреляционному алгоритму по вычислительной сложности.

На рис. 7 представлены примеры комбинированного изображения, состоящего из телевизионного изображения и изображения цифровой карты местности. Изображение на рис. 7, а построено для исходных навигационных параметров, а изображение на рис. 7, б – с учетом найденных параметров геометрических преобразований изображений.

Заключение

Разработанный алгоритм синтеза комбинированного изображения, включающий выделение границ на основе нечеткой кластеризации и оценивание параметров геометрических преобразований на основе свойств преобразования Фурье, подтвердил свою работоспособность в ходе экспериментальных исследований. Так при грубом оценивании параметров геометрических преобразований изображений достигается ускорение работы в 20 раз, а при приближении точности оценивания к максимуму, доступной при заданном разрешении съемки, производительность увеличивается в тысячи раз по сравнению с методом полного перебора и в 3 раза по сравнению с корреляционным подходом.

Таблица 1. Сравнение времени работы при шаге перебора углов равном 1°

	Всего, с.	Рендеринг карты, с.	Выделение границ, с.	Оценивание рассогласования, с.
Полный перебор навигационных параметров	647,9	644,6	0,035	3,3
С использованием расширенного изображения местности	24,1	23,9		0,16
С использованием преобразования Фурье	31,4			7,5

Таблица 2. Сравнение времени работы при шаге перебора углов равном 0,125°

	Всего, с.	Рендеринг карты, с.	Выделение границ, с.	Оценивание рассогласования, с.
С использованием расширенного изображения местности	105	23,9	0,035	81,1
С использованием преобразования Фурье	32,9			9



а)



б)

Рис. 7. Комбинированные изображения: исходные параметры (а), результат устранения рассогласования (б)

Работа выполнена за счет средств гранта Президента РФ для государственной поддержки молодых российских ученых МК-3732.2015.10.

Литература

1. Alpatov B., Babayan P., Khosenko M. Image Synthesis Using Searching and Tracking Techniques in Combined Vision Systems // Proceedings of 4th Mediterranean Conference on Embedded Computing (MECO). – 2015. – P. 147-150.
2. Мишин А.Ю., Кирюшин Е.Ю. и др. Малогабаритная комплексная навигационная система на микромеханических датчиках // Труды МАИ. – 2013. – № 70. – С. 1-21.
3. Alpatov B.A., Strotov V.V. An estimation algorithm of the multispectral image geometric transformation parameters based on multiple reference area tracking // Proceedings of the SPIE, 2013. – Vol. 8713. – 8 p.
4. Szeliski R. Computer Vision: Algorithms and Applications. – London: Springer-Verlag, 2011. – 812 p.
5. 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.
6. Визильтер Ю.В., Желтов С.Ю. и др. Обработка и анализ изображений в задачах машинного зрения. Курс лекций и практических занятий. – М.: Физматкнига, 2010. – 672 с.
7. Dunn J.C. A Fuzzy Relative of the ISODATA Process and Its Use in Detecting Compact Well-Separated Clusters // Journal of Cybernetics. – 1973. – Vol. 3. – P. 32-57.
8. MacQueen J.B. Some Methods for Classification and Analysis of Multivariate Observations // Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability. – 1967. – Vol. 1. – P. 281-297.
9. Cormen T.H., Leiserson C.E., Rivest R.L., Stein C. Introduction to Algorithms, 2nd ed. – MIT Press, 2001. – 1292 p.

УДК 681.513.3

АЛГОРИТМ ОБНАРУЖЕНИЯ НАЗЕМНЫХ ПРЕПЯТСТВИЙ ПРИ МАЛОВЫСОТНЫХ ПОЛЕТАХ

*Балашов О.Е., к.т.н., доцент кафедры АИТУ Рязанского государственного радиотехнического университета,
e-mail: aitu@rsreu.ru.*

OBSTACLE DETECTION ALGORITHM AT LOW-ALTITUDE FLIGHT

Balashov O.E.

The problem of detection of ground obstacle at low-altitude flight is described in this paper. It presents an algorithm to measure the distance between the aircraft and the obstacle using one video sensor. Obstacle detection is based on the analyses of coordinates of feature points at the image sequence. Points lying above the earth's surface considered to be an obstacle and calculate the times to collision to them.

Keywords: obstacle detection, time to collision, measurement of distance to objects.

Ключевые слова: оптико-механическая система, обнаружение препятствий, расстояние до объектов.

Введение. Постановка задачи

При выполнении полетов над населенными пунктами на малой высоте большую опасность для летательных аппаратов с точки зрения столкновения представляют высотные сооружения (опоры мостов, высотные здания, столбы и т.п.). Для уменьшения опасности полетов целесообразно использовать системы помощи, которые способны заблаговременно предупредить пилота о приближении летательного аппарата к высотному препятствию. Системы предупреждения столкновений могут строиться на базе различных датчиков, выдающих информацию об окружающем пространстве. В данной статье рассматривается подход, позволяющий оценить расстояние до наземных объектов с помощью обработки последовательности видеок кадров, сформированных одним видеодатчиком (ВД). На основании информации о дальности до высотных сооружений система предупреждения столкновений способна принимать решение об опасности полета и может выдавать пилоту предупреждения о препятствии и советы по маневрированию для предотвращения столкновения.

Обнаружение и оценка параметров препятствий основаны на анализе координат точек наблюдаемых объектов в последовательности видеок кадров. Для обнаружения и выделения точек объектов в последовательности изображений необходимо чтобы их изображения обладали какими-либо отличительными признаками, то есть точки должны быть ключевыми. Существуют многочисленные точечные операторы (детекторы ключевых точек), предназначенные для выделения особых (характерных) точек на изображении. Для большинства возможных типов препятствий изображения, содержащие границы высотных сооружений и фоновых объектов, будут иметь угловые точки. Поэтому для обнаружения препятствий удобно использовать детекторы углов, например, детектор Харриса [1]. Благодаря анализу то-

Рассматривается один из возможных способов оценки расстояния между летательным аппаратом, являющимся носителем оптико-механической системы, и наземным высотным сооружениями. Приводится алгоритм вычисления пространственных координат объектов с использованием одного видеодатчика.

чек, лежащих на краях изображений препятствий, возможно более точное оценивание геометрических параметров препятствий, по сравнению с вариантом, когда анализируются ключевые точки, лежащие только внутри (не на границе) изображения высотного сооружения.

Для обнаружения препятствий носитель вместе с ВД должен совершать поступательное движение. В результате будем иметь последовательность изображений, содержащих изображения объектов (препятствий) с разных ракурсов (рис. 1).

Системы предупреждения столкновений удобно строить на базе бортовых оптико-механических систем (ОМС), предназначенных для сопровождения объектов, мониторинга окружающей обстановки [2-4]. Видеодатчик в ОМС обычно крепится на внутренней рамке карданного подвеса, устанавливаемого в носовой части носителя (летательного аппарата) (рис. 2). Будем считать, что в моменты формирования видеок кадров система предупреждения столкновений обладает информацией о скорости движения носителя, его координатах и ориентации в пространстве. Также будем считать известными углы поворота рамок кардана (угловые координаты оптической оси ВД) θ , φ , ψ относительно носителя. Данная информация может поступать в рассматриваемую систему с бортовой навигационной аппаратуры и датчиков углов поворота рамок кардана.

На рис. 2, а изображен вариант ОМС с трехосным карданным подвесом. Для обнаружения препятствий также могут использоваться ОМС с двухосными карданами. Вращение ВД относительно носителя осуществляется оператором или какими-либо бортовыми системами, например, системой автоматического сопровождения объектов по видеоданным [2, 5, 6].

Задача обнаружения препятствий перед летатель-

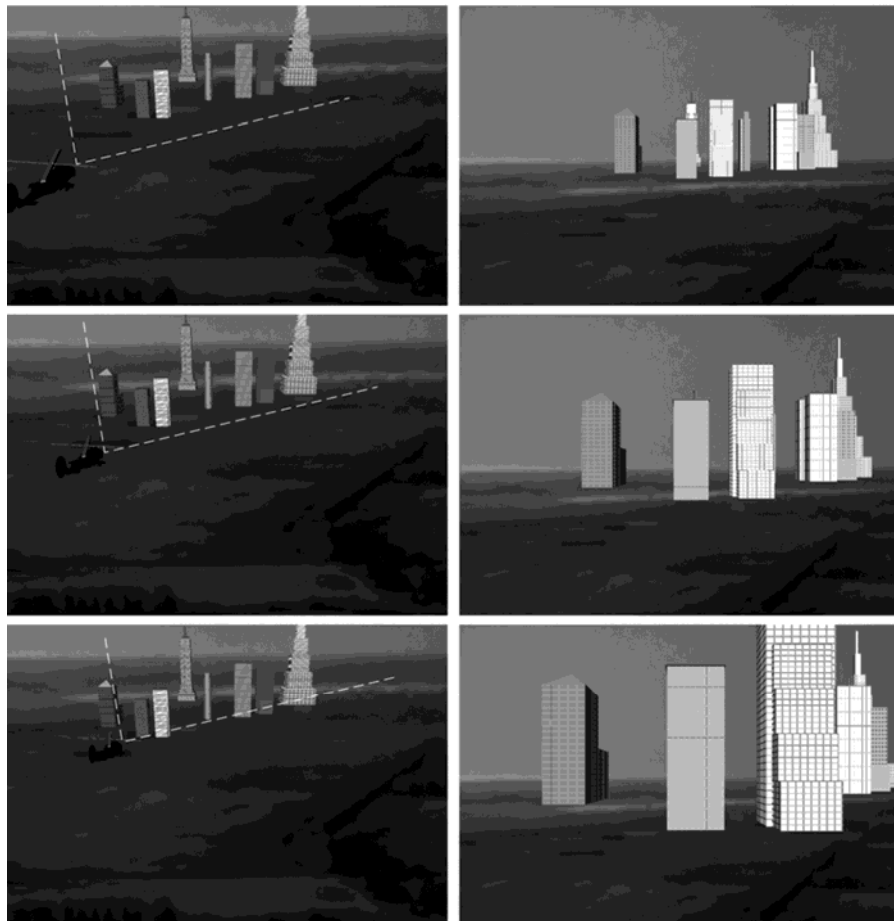


Рис. 1. Пример движения носителя и обрабатываемых изображений

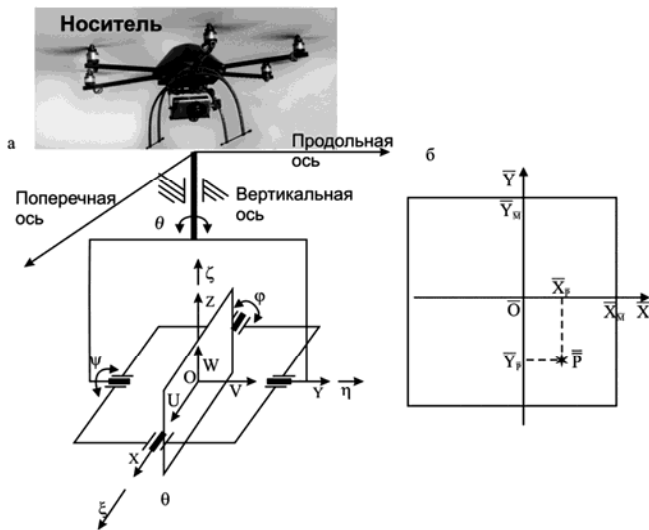


Рис. 2. Крепление ВД на носителе ОМС

ным аппаратом может решаться как при вращающемся ВД, так и при отсутствии вращения рамок кардана. В некоторых случаях ВД может быть жестко закреплен на носителе, тогда углы поворота рамок кардана θ , φ , ψ имеют постоянные значения.

Алгоритм обнаружения и измерения параметров препятствий

При выполнении маловысотных полетов обнаружение наземных препятствий возможно путем анализа в последовательности изображений координат пикселей,

являющихся изображением объектов. Для слеживания за точками объекта в последовательности кадров удобно использовать ключевые точки изображения, выделяемые детектором углов. На рис. 3 показан пример анализируемого изображения, на котором показаны ключевые точки и смещение точек между двумя кадрами.



Рис. 3. Анализируемое изображение

Рассмотрим некоторое высотное сооружение, одну его из точек, выделяемую на изображении как ключевую, обозначим P . На рис. 2, а показан вариант трехосного карданного подвеса, расположенного на носителе, с изображением соответствующих систем координат. Положение системы координат $OUVW$, жестко связанной с ВД, относительно носителя определяется углами поворота рамок карданного подвеса (θ – угол азимута, φ –

угол места, ψ – угол крена). Направление оптической оси ВД совпадает с осью OV . Система координат $O\xi\eta\zeta$ неподвижно ориентирована относительно поверхности земли так, что плоскость $O\xi\eta$ параллельна плоскости горизонта, а ось $O\xi$ образует правую систему координат с плоскостью $O\xi\eta$. Положение системы координат носителя $OXYZ$, образуемой строительными осями носителя, относительно $O\xi\eta\zeta$ определяется углами качки (α – угол рысканья, β – тангажа, γ – крена). При отсутствии качки носителя и равенстве нулю углов поворота рамок карданного подвеса θ , ϕ и ψ все указанные системы координат совпадают. Пусть текущим значениям α , β , γ , θ , ϕ , ψ соответствует плоское изображение, получаемое с ВД и представленное на рис. 2, б, где отрезок \overline{OP} определяет положение анализируемой точки P .

В качестве модели изображения с угловым полем зрения $2X_M \times 2Y_M$ примем изображение $ABCD$ (рис. 4) размером $2N_{XM} \times 2N_{YM}$, формируемое на плоскости, перпендикулярной оптической оси видеодатчика OV и расположенной на расстоянии R от оптического центра ВД [7]. Размер изображения в пикселях равен $2N_{XM} \times 2N_{YM}$. При этом предполагается, что любая произвольная точка изображения $ABCD$ воспринимает лишь световой поток, проходящий через заданную точку и центр O оптической системы. Пусть анализируемой точки объекта на изображении $ABCD$ соответствует пиксель P_i с координатами NXP, NYP . Направление на анализируемую точку в момент времени, соответствующий формированию i -го видеокadra, задается вектором \overline{OP}_i . Начало \overline{OP}_i совпадает с началом системы координат $OUVW$.

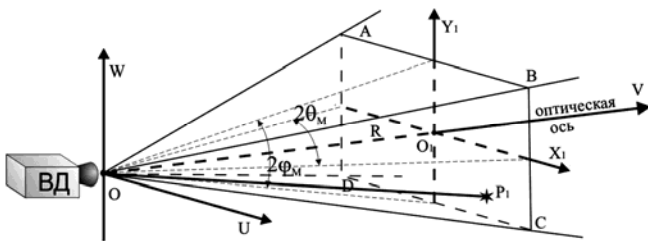


Рис. 4. Модель изображения, формируемого видеодатчиком

Найдем направление на объект \overline{OP}_i в системе координат $O\xi\eta\zeta$ в момент формирования первого видеокadra. Рассмотрим систему координат модели изображения $O_1X_1Y_1$. Координаты точки P_1 и вектора \overline{OP}_1 в системе координат $OUVW$ легко вычислить по координатам изображения точки $P_1(NXP, NYP)$, вычисляемым в результате обработки текущего видеокadra в системе координат изображения $O_1X_1Y_1$ (рис. 2, б). В результате получим, что направление на анализируемую точку P на i -м кадре в системе координат $OUVW$ описывается следующим вектором [2]

$$\overline{OP}_i = \begin{bmatrix} \frac{N_{XP}}{N_{XM}} \operatorname{tg} \theta_M & 1 & \frac{N_{YP}}{N_{YM}} \operatorname{tg} \phi_M \end{bmatrix} \quad (1)$$

Чтобы получить координаты вектора \overline{OP}_i в системе координат $O\xi\eta\zeta$, необходимо произвести преобразование координат с помощью матрицы направляющих косинусов $T'(\theta, \phi, \psi)$, переводящей координаты вектора из системы координат $OUVW$ в $OXYZ$, и $T(\alpha, \beta, \gamma)$, переводящей координаты из $OXYZ$ в $O\xi\eta\zeta$ [8]. Матрица $T(\theta, \phi, \psi)$ имеет вид

$$T(\theta, \phi, \psi) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\phi) & \sin(\phi) \\ 0 & -\sin(\phi) & \cos(\phi) \end{bmatrix} \times \begin{bmatrix} \cos(\psi) & 0 & -\sin(\psi) \\ 0 & 1 & 0 \\ \sin(\psi) & 0 & \cos(\psi) \end{bmatrix} \times \begin{bmatrix} \cos(\theta) & \sin(\theta) & 0 \\ -\sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (2)$$

а $T(\alpha, \beta, \gamma)$ может быть получена из матрицы $T(\theta, \phi, \psi)$ заменой θ на α , ϕ на β , ψ на γ . В результате координаты вектора \overline{OP}_i в системе координат $O\xi\eta\zeta$ описываются следующим соотношением

$$\overline{p}_i = T(\alpha, \beta, \gamma) \cdot T(\theta, \phi, \psi) \cdot \overline{OP}_i \forall i, \quad (3)$$

Дальность до анализируемой точки P можно рассчитать на основании направлений \overline{p}_i , полученных из разных точек пространства за счет движения носителя. Таким образом, для вычисления дальности необходимо решить задачу триангуляции.

Рассмотрим движение носителя в трехмерном пространстве. Введем неподвижную систему координат $O'\xi'\eta'\zeta'$, оси которой сонаправлены с осями системы координат $O\xi\eta\zeta$, а плоскость $O'\xi'\eta'$ совпадает с плоскостью земли и параллельна плоскости $O\xi\eta$. Началом системы координат $O'\xi'\eta'\zeta'$ выберем проекцию точки O на плоскость земли в момент формирования первого видеокadra. На рис. 5 показано положение носителя (точки $T_1(\xi_1, \eta_1, \zeta_1)$, $T_2(\xi_2, \eta_2, \zeta_2)$, $T_3(\xi_3, \eta_3, \zeta_3)$) и направления на анализируемую точку пространства $P(\xi, \eta, \zeta)$ в системе координат $O'\xi'\eta'\zeta'$. Точка T_1 соответствует положению носителя в момент формирования первого видеокadra, точка T_2 – в момент формирования второго видеокadra и т.д. Точки T'_1, T'_2, T'_3 являются проекциями точек T_1, T_2, T_3 на плоскость $O'\xi'\eta'$. Так как проекция точки T_1 совпадает с началом системы координат $O'\xi'\eta'\zeta'$, то координаты точки T_1 равны $(0, 0, \zeta_1)$, ζ_1 – высота полета носителя в момент времени, соответствующий первому видеокadру. На рис. 5 точка P лежит в плоскости $O'\xi'\eta'$, поэтому она совпадает с проекцией P' на плоскость $O'\xi'\eta'$ (это не влияет на ход рассуждений).

Процедуру вычисления дальности до анализируемой точки P рассмотрим в момент формирования третьего видеокadra (точка T_3). Для решения поставленной задачи необходима информация о направлении на анализируемую точку в любые два момента времени. Для оценки дальности на текущем видеокadre целесообразно

выбрать направление, полученное в текущий момент времени \overline{p}_3 , и направление, вычисленное на одном из предшествующих кадров, например, \overline{p}_2 . В этом случае направления на анализируемую точку задаются соответственно прямыми T_3P , T_2P . Так как в цифровых системах всегда присутствуют ошибки дискретизации, то прямые T_3P , T_2P , полученные в результате обработки изображений и навигационной информации о движении носителя, скорее всего, не будут пересекаться, хоть и описывают направление на одну точку пространства. В этом случае точкой P можно считать такую точку пространства, к которой обе прямые сближаются наиболее близко. Найти подобную точку можно решив задачу минимизации критерия, описывающего сумму квадратов расстояний искомой точки P до прямых.

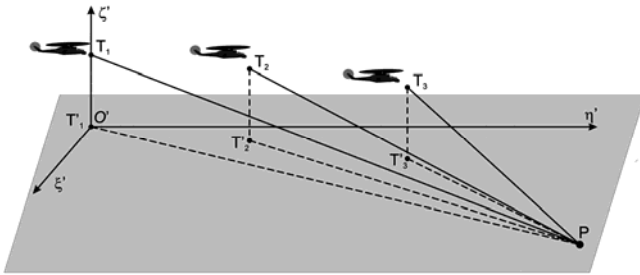


Рис. 5. Определение дальности

В данной статье предлагается другой способ. Сначала найдем на плоскости $O'\xi'\eta'$ точку пересечения проекций прямых T_3P , T_2P . Точкой пересечения проекций является точка P' . Решение может быть найдено, если проекции не параллельны. Параллельные проекции могут получиться из-за ошибок дискретизации при малых изменениях направления на объект в последовательности видеок кадров, при большой удаленности анализируемой точки P от ВД, при движении строго на точку P . В первых двух случаях для решения задачи поиска дальности до точки P , а также в некоторых случаях для повышения точности решения следует выбирать текущее направление на объект и направление, полученное несколько кадров назад (например, направление \overline{p}_1). Чем больше времени прошло между обрабатываемой парой видеок кадров, тем меньше будет влияние ошибок дискретизации изображений. Это достигается за счет усреднения измеряемых угловых координат направлений на анализируемую точку. В третьем случае для оценки дальности до точки P следует использовать проекции направлений на вертикальную плоскость $O'\eta'\zeta'$.

Для отыскания координат точки P' составим векторные параметрические уравнения прямых T_3P , T_1P . Прямая T_3P проходит через точку $T_3(\xi_3, \eta_3, \zeta_3)$ и задается направляющим вектором $\overline{p}_3(\xi_{P_3}, \eta_{P_3}, \zeta_{P_3})$. Прямая T_1P проходит через точку $T_1(\xi_1, \eta_1, \zeta_1)$ и задается вектором $\overline{p}_1(\xi_{P_1}, \eta_{P_1}, \zeta_{P_1})$. Тогда уравнения прямых T_3P , T_1P , обозначим их соответственно l_3 , l_1 , описываются следующими выражениями

$$l_3 = T_3 + t_3 \overline{p}_3; \quad l_1 = T_1 + t_1 \overline{p}_1, \quad (4)$$

где t_1 , t_3 – параметры прямых. Для нахождения координат

точки P' следует составить систему уравнений, приравняв уравнения прямых l_1 , l_3 по координатам ξ и η . В результате получим систему из двух уравнений с двумя неизвестными t_1 , t_3 .

$$\begin{cases} \xi_1 + t_1 \xi_{P_1} = \xi_3 + t_3 \xi_{P_3}; \\ \eta_1 + t_1 \eta_{P_1} = \eta_3 + t_3 \eta_{P_3}. \end{cases} \quad (5)$$

В силу не параллельности направлений на анализируемую точку существует единственное решение t'_1 , t'_3 системы (5). Получив решение (5) координаты точки $P(\xi\eta\zeta)$ можно найти по выражению

$$P = T_3 + t'_3 \overline{p}_3 \quad \text{или} \quad P = T_1 + t'_1 \overline{p}_1, \quad (6)$$

По координатам точки P легко вычислить дальность D_P до анализируемой точки по следующему выражению

$$D_P = \sqrt{\xi^2 + \eta^2 + \zeta^2}. \quad (7)$$

Координата ζ точки P позволяет судить о высоте той части объекта, которой принадлежит анализируемая точка. Таким образом, проанализировав в последовательности видеок кадров направления на все точки пространства, попадающие в поле зрения ВД, можно судить о наличии высотных объектов. Высотой объекта считается значение высоты точки, принадлежащей объекту и имеющей наибольшую координату ζ среди всех точек объекта. Для решения задачи сопоставления точек с объектами следует воспользоваться методами кластерного анализа [9]. Точки с ненулевой высотой следует группировать на основании пространственных координат. Точки с нулевой высотой принадлежат лежащим на поверхности земли объектам.

Пример работы алгоритма обнаружения наземных препятствий

Для иллюстрации работы алгоритма рассмотрим случай когда летательный аппарат совершает полет на малой высоте с постоянной скоростью по прямой траектории. Будем считать известными в моменты формирования видеок кадров координаты и скорость носителя, а также углы поворота рамок кардана. Для простоты описания положим в данном примере, что углы поворота кардана постоянны и равны нулю. Пусть перед летательным аппаратом имеются препятствия в виде трех высотных зданий. На рис. 6, а показан один из видеок кадров обрабатываемой тестовой видеопоследовательности. На рис. 6, б показаны ключевые точки, выделенные на изображении с помощью детектора особых точек (детектора Харриса). На рис. 6, в изображено перемещение ключевых точек между первым и вторым анализируемыми видеок кадрами. На рис. 6, г показаны результаты обнаружения высотных препятствий и ключевые точки изображения, для которых оценка высоты больше нуля.

Заключение

Описанный алгоритм вычисляет дальность до препятствий и их высоту над поверхностью земли путем обработки последовательности видеок кадров, сформированных одним видеодатчиком. Для применения рассмотренного алгоритма требуется постоянное поступательное движение видеодатчика. Результатом работы

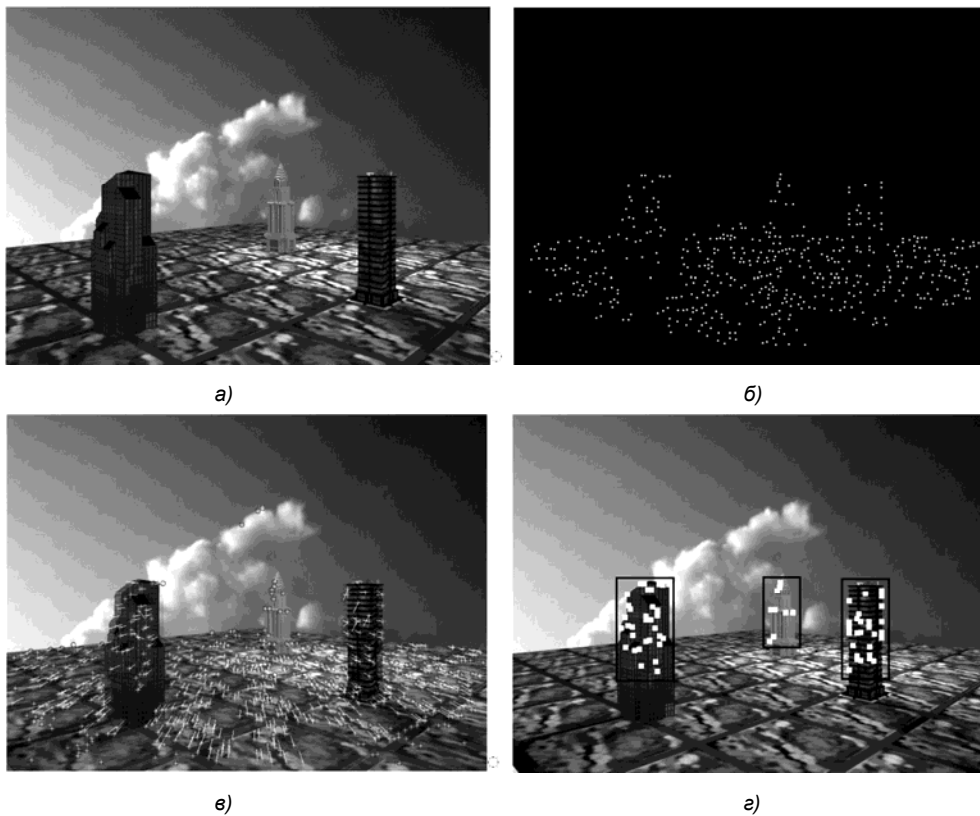


Рис. 6. Этапы алгоритма обнаружения препятствий

алгоритма является информация о направлении на объект, дальность до точек объекта, пространственные координаты препятствия.

Точность оценки параметров высотных объектов зависит от скорости носителя, дальности до препятствий, их расположения на изображении относительно центра изображения. Данные параметры определяют скорость изменения угловых координат точек объекта при движении носителя. Для повышения точности оценки координат точек объекта следует выбирать более длинную предысторию видеокadres. Также на точность вычисления координат ключевых точек сильное влияние оказывают ошибки дискретизации изображений, датчиков углов поворота рамок кардана, датчиков высоты и ориентации носителя, а также ошибки сопоставления ключевых точек на анализируемых видеокadres. Анализ точности измерения координат точек объекта показывает, что высотные сооружения в виде зданий могут быть обнаружены на расстоянии до 1 км.

Исследования выполнены при финансовом содействии гранта Президента РФ МК-433.2014.10.

Литература

1. Визильтер Ю.В., Желтов С.Ю., Бондаренко А.В., Ососков М.В., Моржин А.В. Обработка и анализ изображений в задачах машинного зрения: Курс лекций и практических занятий. – М.: Физматкнига, 2010. – 672 с.
2. 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.: ил.

вождения объектов. Обработка изображений и управление. – М.: Радиотехника, 2008. – 176 с.: ил.

3. Балашов О.Е., Степашкин А.И. Нашлемная система обзора и целеуказания // Вестник Рязанского государственного радиотехнического университета. Вып. 4 (38). - Рязань, 2011. – С. 40-44.

4. Алпатов Б.А., Балашов О.Е., Степашкин А.И. Прогнозирование угловых координат движущихся объектов в бортовых оптико-механических системах // Информационно-управляющие системы. 2011. № 5. С. 2 – 7

5. Алпатов Б.А., Балашов О.Е., Степашкин А.И., Трофимов Д.В. Алгоритм измерения угловых координат линии визирования оператора // Информационно-управляющие системы. 2012. № 3. С. 18–21.

6. Алпатов Б.А., Балашов О.Е., Степашкин А.И., Трофимов Д.В. Алгоритм вычисления угловых координат линии визирования оператора в нашлемной системе позиционирования // Информационно-управляющие системы. 2012. № 6 (61). С. 7-11.

7. Алпатов Б.А., Балашов О.Е., Степашкин А.И. Некоторые вопросы построения видеокomпьютерных следящих систем // Проблемы математического моделирования и обработки информации в научных исследованиях: Сб. науч. тр. Рязань: РГРТА, 2003. С. 16–25.

8. Бесекерский В.А., Фабрикант Е.А. Динамический синтез систем гироскопической стабилизации. Л.: Судостроение, 1968. 351 с.

9. Калинина В. Н., Соловьев В. И. Введение в многомерный статистический анализ: Учебное пособие / ГУУ – М., 2003. – 66 с.

ОЦЕНКА ВЛИЯНИЯ ШУМА НА ТЕЛЕВИЗИОННОЕ ИЗОБРАЖЕНИЕ СИСТЕМЫ ПОЗИЦИОНИРОВАНИЯ В ПОМЕЩЕНИИ

Тюкин А.Л., аспирант кафедры динамики электронных систем, Ярославский государственный университет им. П.Г. Демидова, e-mail: tyukin.alex@gmail.com;

Приоров А.Л., д.т.н., доцент, доцент кафедры динамики электронных систем, ЯрГУ им. П.Г. Демидова, e-mail: andcat@yandex.ru.

THE ESTIMATION OF THE NOISE INFLUENCE ON THE TELEVISION IMAGE OF THE INDOOR POSITIONING SYSTEM

Tyukin A.L., Priorov A.L.

The article describes indoor positioning algorithm based on digital image processing. The image goes from industrial television system. Reference points are color beacons, which have a constant coordinates indoor. The paper describes the carried out estimation of the algorithm resistance to different types of noise influence on input television image.

Keywords: industrial television system, digital image processing, mobile robotic platform, indoor positioning, color beacons, estimation of the noise influence.

Ключевые слова: прикладное телевидение, цифровая обработка изображений, мобильная роботизированная платформа, позиционирование в помещении, цветовой маяк, оценка влияния шума.

Введение

В настоящее время для решения вопросов позиционирования в помещении создано больше количество систем, работающих на основе разных принципов и с использованием различных типов опорных точек. Но универсального решения, как в глобальной спутниковой радионавигации, пока не найдено. Причиной тому является сложность точного позиционирования в помещении с использованием радиоканала. Это связано с переотражениями радиосигналов и наложением их друг на друга. Установлено, что использование радиоканала для позиционирования в помещении имеет множество проблем [1–4].

С другой стороны, видимый диапазон электромагнитного спектра удобен для практического использования в зоне прямой видимости. Он позволяет четко ориентироваться в пространстве, обладает неплохой помехозащищенностью [5, 6]. Кроме того, для работы в данном диапазоне можно обойтись простой и недорогой цветной телекамерой. В такой системе для ориентации могут быть использованы специальные маяки с цветовой кодировкой [7].

Таким образом, мобильная роботизированная платформа (МРП) с установленной на ней телекамерой может ориентироваться в помещении, в котором находятся цветные маяки известного размера с заданными координатами. Такая монокулярная система прикладного телевидения может оценивать относительный размер маяка на входящем кадре путем цифровой обработки телевизионного изображения и расстояние до него [8, 9]. При нахождении в кадре минимум двух маяков возможно определение местоположения самой телекамеры, а, следовательно, и МРП.

Рассматривается алгоритм позиционирования в помещении, основанный на цифровой обработке изображения, поступающего из системы прикладного телевидения. В качестве опорных точек предложено использовать цветные маяки, статично распрямленные в помещении с заданными координатами. Проводится оценка устойчивости алгоритма к влиянию различных типов шумов на входящее телевизионное изображение.

Во многих случаях расчётный модуль, который проводит обработку входящей видеопоследовательности, целесообразно вынести за пределы МРП, что позволит упростить устройство самой платформы и заметно снизить её энергопотребление, которое часто является достаточно критическим параметром. Для обмена данными с расчётным модулем в этом случае необходим дополнительный приёмо-передающий модуль.

В результате передачи данных по каналу связи на полезный сигнал могут накладываться помехи, что, в итоге, может привести к появлению шума на полученном изображении. Кроме того, шумы на изображении в данном случае могут возникать при низкой освещённости сцены (за счёт возрастания параметров передаточной функции телекамеры) [10, 11], и в связи с внутренними шумами телекамеры (например, при высоком уровне теневого тока фотосенсора) [12].

Целью работы является разработка алгоритма позиционирования в помещении по маякам с цветовой кодировкой на основе цифровой обработки изображения, получаемого от системы прикладного телевидения. Рассматривается устойчивость разработанного алгоритма к шумам различного типа, существенно влияющим на качество телевизионного изображения.

Краткое описание работы алгоритма позиционирования

Работа алгоритма позиционирования в помещении основана на цифровой обработке изображения, поступающего из системы прикладного телевидения. Данная

система включает в себя простую цветную телекамеру, установленную на МРП, расчётный модуль, проводящий обработку видеоизображения с телекамеры, и приёмопередающий блок, обеспечивающий обмен данными между МРП и расчётным модулем. Для ориентации в помещении в качестве опорных точек предложено использовать специальные маяки с цветовой кодировкой (рис. 1). Преимущество таких маяков в сравнении с другими [1, 2] состоит в том, что они просты в изготовлении, экономичны и не требуют источников питания, что позволяет им оставаться работоспособными достаточно долгое время.

Обнаружение маяков и оценка их относительного размера на входном телевизионном изображении позволяют системе оценить их относительные координаты, а зная априорно заданные абсолютные координаты цветных маяков, можно с заданной точностью определить местоположение телекамеры (абсолютные координаты телекамеры) [13]. Структурная схема соответствующего алгоритма представлена на рис. 2.

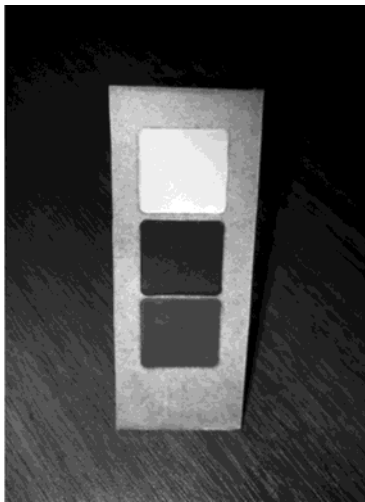


Рис. 1. Маяк с цветовой кодировкой

Для решения задачи распознавания на маяки накладываются три основных ограничения [14, 15]:

1. Все три цвета на маяке должны быть визуально хорошо различимы.
2. Центры цветных областей должны располагаться на одной прямой и должны быть равноудалены друг от друга.
3. Поверхность маяков должна быть матовой (не должна бликовать).

Первое ограничение связано с использованием фильтрации по цвету каналов HSV изображения: для оттенка (Hue) применяется гауссов колокол с варьируемой дисперсией [16]:

$$f(x) = 255 \exp\left(-\frac{(x-m)^2}{2D}\right), \quad (1)$$

а для насыщенности и яркости – логистический сигмоид с варьируемой кривизной и смещением:

$$f(x) = \frac{255}{1 + \exp(-k(x - m + \Delta))}, \quad (2)$$

где D – дисперсия значения оттенка пикселя, m – значение, соответственно, оттенка, насыщенности или яркости одной из трёх цветных областей маяка, x – значение анализируемого пикселя, k – кривизна логистического сигмоида, Δ – смещение.

Второе ограничение отбраковывает сторонние цветные области, соответствующие по цвету одной из трёх областей маяка [13]. Третье ограничение позволяет избежать искажения цвета области из-за возможного возникновения бликов при ярком освещении.

После применения функций (1) и (2) получаются три изображения (для H , S и V каналов), которые попиксельно усредняются для получения одного изображения для отдельного цвета [13]. Усреднение проводилось с помощью среднего геометрического взвешенного:

$$\bar{x} = x_H^{w_H} \cdot x_S^{w_S} \cdot x_V^{w_V}, \quad (3)$$

где $w_H + w_S + w_V = 1$, w – вес. Данный выбор сделан в связи с тем, что при таком способе усреднения учитываются особенности цвета, а именно – оттенок берется с большим весом, чем насыщенность и яркость ($w_H > w_S, w_V$), а последние – с весами, равными: $w_S = 2 w_V$. Это сделано потому, что цвет – главный критерий работы алгоритма детектирования, который в системе HSV кодируется двумя значениями цветности: оттенком H и насыщенностью S . Оттенок характеризует цвет области, насыщенность – чистоту цвета. Яркость характеризует интенсивность света и связана с параметрами источника света, а не свойствами поверхности.

Для полученных областей вычисляются геометрические центры. После определения всех возможных центров цветных областей маяков, находящихся в кадре, вычисляются длины и углы наклона (относительно вертикали) векторов, соединяющих центры залитых областей (рис. 3).

Относительные координаты маяков определяются, исходя из их относительного размера во входном изображении. В этом случае камера имеет нулевые координаты на относительной карте [17]. Для расчёта относительных координат маяка используются формулы:

$$d = \frac{R/2}{\operatorname{tg}(\psi/2)} \cdot \frac{h}{h_{\text{пикс}}}, \quad (4)$$

$$x = x_{\text{пикс}} \frac{h}{h_{\text{пикс}}}, \quad z = z_{\text{пикс}} \frac{h}{h_{\text{пикс}}}, \quad (5)$$

где d – расстояние от камеры до плоскости, в которой расположен маяк (координата y), m , R – разрешение камеры по вертикали, пиксель; ψ – апертура объектива по вертикали, градус; h – реальная высота маяка, $h_{\text{пикс}}$ – относительная высота маяка на изображении, пиксель; $x_{\text{пикс}}$ и $z_{\text{пикс}}$ – смещение центра маяка от оптической оси телекамеры по горизонтали и по вертикали соответственно,



Рис. 2. Структурная схема алгоритма

пиксель; x и z – соответствующие координаты маяка.

Для получения местоположения МРП строится абсолютная карта. Переход к абсолютным координатам телекамеры обеспечивается через использование трехмерного преобразования относительных координат маяков. Абсолютная карта – это изображение с планом помещения или пространства, по которому движется автономная МРП. К ней жестко «привязывается» декартова система координат [16].

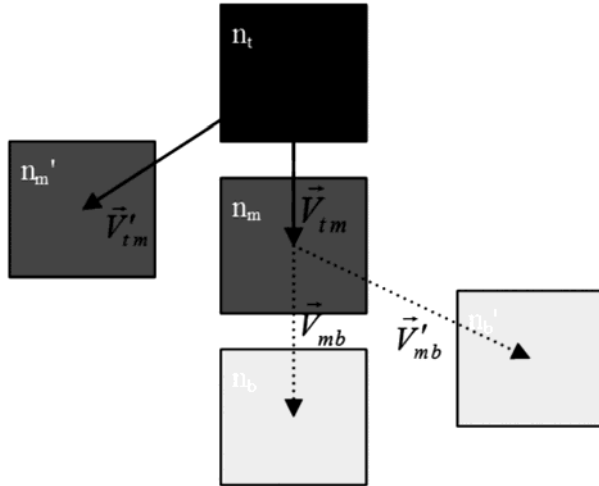


Рис. 3. Всевозможные комбинации векторов

В алгоритме позиционирования используются два метода преобразования координат: метод трехмерных аффинных преобразований (поворот координатной плоскости, масштабирование и параллельный перенос) и метод градиентного спуска (параллельный перенос, поворот по достижению минимума СКО, масштабирование по достижению минимума СКО) [13].

Изучение влияния шума на работу алгоритма

На работу алгоритма позиционирования могут влиять шумы различной природы. Они могут возникать на разных этапах работы системы. Степень зашумленности зависит от интенсивности и вида помех [18]. При проектировании системы позиционирования необходимо учесть возможное влияние подобных факторов.

Для надежного функционирования алгоритма позиционирования важными параметрами являются его устойчивость к шуму и величина ошибки измерений, вызванная шумами различной интенсивности и природы. На сегодняшний день наиболее используемым показателем оценки качества является пиковое отношение сигнал/шум (для изображений с 256 возможными значениями яркости):

$$ПОСШ = 10 \lg \frac{255^2}{СКО}, \quad (6)$$

$$СКО = \frac{1}{N} \sum_{i=1}^N (x_i - y_i)^2, \quad (7)$$

где N – число пикселей изображения, x_i и y_i – значения пикселей исходного и зашумленного изображений соответственно, а СКО – среднеквадратичная ошибка [19].

Знание причины возникновения шумов, визуальный анализ зашумленных кадров помогают в построении модели шума и оценке его статистических характеристик

[20, 21]. Для исследования влияния шума на предложенный алгоритм позиционирования исследованы следующие модели шумов:

– Аддитивный белый гауссовский шум (АБГШ). Такая модель может быть описана соотношением зашумленного изображения: $g(x, y) = f(x, y) + \eta(x, y)$, где $f(x, y)$ – исходное изображение, $\eta(x, y)$ – независимый от сигнала АБГШ.

– Мультипликативный шум. Такой тип шума описывается выражением $g(x, y) = f(x, y) \cdot \eta(x, y)$.

– Шум «соль и перец» представляется выражением $g(x, y) = (1 - p)f(x, y) + p \cdot i(x, y)$, где $i(x, y)$ – модель импульсного шума, p – бинарный параметр со значениями 1 и 0, который и отражает максимальные и минимальные значения пикселя – «соль» и «перец».

При расположении двух маяков в объективе телекамеры проведены исследования зависимости величины дисперсии абсолютных координат камеры от ПОСШ. Для этого записывалась видеопоследовательность длительностью 180 с, на которой присутствовали два маяка. При этом изображение оставалось статичным на всем протяжении эксперимента (ни маяки, ни камера не двигались), а освещенность – постоянной. На первом этапе эксперимента для преобразования относительных координат маяков в абсолютные координаты камеры применялся метод аффинных преобразований.

После формирования видеозаписи на неё накладывался шум различной природы и интенсивности. Снимались значения ПОСШ и дисперсии абсолютных координат камеры. По полученным данным строились графики (рис. 4).

Видно, что при уровне ПОСШ больше 16 дБ все три вида шума оказывают на алгоритм позиционирования примерно одинаковое влияние. Установлено, что граница работоспособности алгоритма при наложении гауссова шума равна 16 дБ. При меньшем значении ПОСШ алгоритм не способен распознать маяк.

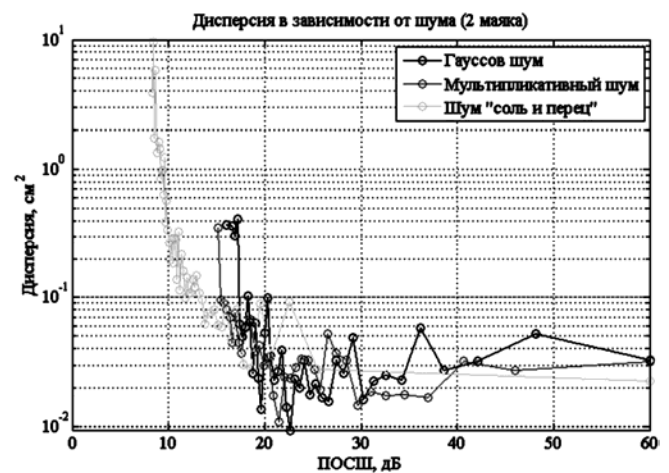


Рис. 4. Зависимость дисперсии абсолютных координат камеры от шума при использовании метода аффинных преобразований и детектировании 2 маяков

При воздействии мультипликативного шума алгоритм дольше сохраняет способность распознавания маяка. Граница работоспособности смещается до уровня 15,1 дБ. Однако при таком ПОСШ дисперсия резко возрастает (примерно на порядок).

К шуму «соль и перец» алгоритм оказался наиболее устойчив. Граница работоспособности составила 8,3 дБ, но и дисперсия при таком ПОСШ максимальна (примерно на два порядка больше средней).

На втором этапе исследования для преобразования координат использовался метод градиентного спуска (рис. 5).

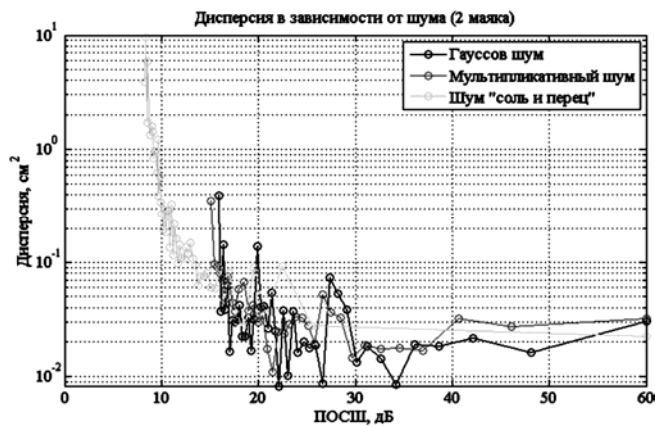


Рис. 5. Зависимость дисперсии абсолютных координат камеры от шума при использовании метода градиентного спуска и детектировании 2 маяков

Также эксперимент проведен и при распознавании трёх маяков. Полученные результаты для двух и трёх маяков с применением различных методов преобразования координат приведены в табл. 1. Для сравнения здесь добавлены результаты эксперимента по зависимости дисперсии координат центра маяка от ПОСШ при распознавании одного маяка в кадре.

Из анализа полученных результатов исследования установлено, что при увеличении количества маяков в кадре ситуация качественно не изменяется и устойчивость к влиянию шума существенно не увеличивается. Следовательно, для повышения устойчивости алгоритма к влиянию шума использовать четыре и более маяков для позиционирования нецелесообразно, поскольку это приведет в большей степени лишь к понижению производительности алгоритма [22].

Выяснено, что алгоритм наименее устойчив к влиянию гауссова шума, и при уровне ПОСШ ниже 16,6 дБ теряет работоспособность. При влиянии мультиплика-

тивного шума границей работоспособности является уровень 15,9 дБ. Наиболее низкий порог работоспособности алгоритм показал при воздействии на него шума «соль и перец» – 9,1 дБ при детектировании трёх маяков. Причём, высокая устойчивость к воздействию шума «соль и перец» сопровождается резким возрастанием дисперсии абсолютных координат камеры. Значение дисперсии на пороге работоспособности более чем на порядок превышает среднее значение дисперсии при различных уровнях ПОСШ.

Сравнение результатов работы алгоритмов преобразования координат показывает, что при использовании метода градиентного спуска система ведет себя более стабильно и показывает низкие значения дисперсии определения координат в сравнении с результатами при использовании метода аффинных преобразований.

Заключение

В работе описан разработанный алгоритм позиционирования в помещении по маякам с цветовой кодировкой. Для работы алгоритма на мобильную роботизированную платформу устанавливается телекамера для фиксации окружающего пространства и распознавания цветowych маяков. Такая система прикладного телевидения оценивает относительный размер маяка на входящей видеопоследовательности, а зная его реальные размеры, определяет расстояние до маяка. Телевизионное изображение для цифровой обработки передается по каналу связи на расчётный модуль. В связи с этим и другими внешними причинами на изображение в процессе работы системы могут накладываться помехи.

В работе проведено исследование влияния шума на работоспособность алгоритма. Установлено, что количество маяков в кадре телевизионного изображения существенно образом не влияет на устойчивость системы к влиянию шумов. Алгоритм показал наибольшую устойчивость к влиянию шума «соль и перец» – порог работоспособности при ПОСШ 9,1 дБ. Наименьшая устойчивость зафиксирована при влиянии гауссова шума – порог ПОСШ 16,6 дБ.

Работа выполнена при финансовой поддержке РФФИ (грант № 15 08-99639-а).

Таблица 1. Границы работоспособности алгоритма

Кол-во маяков	Метод преобразования	Гауссов шум (ПОСШ/дисперсия)	Мультипликативный (ПОСШ/дисперсия)	«Соль и перец» (ПОСШ/дисперсия)
1	-	16 дБ / 0,073 пикс. ²	15,1 дБ / 0,22 пикс. ²	7,8 дБ / 1,2 пикс. ²
2	Метод аффинных преобразований	16 дБ / 0,38 см ²	15,1 дБ / 0,35 см ²	8,3 дБ / 9,6 см ²
2	Метод градиентного спуска	16 дБ / 0,39 см ²	14,9 дБ / 0,064 см ²	8,6 дБ / 1,2 см ²
3	Метод аффинных преобразований	16,3 дБ / 0,34 см ²	15,1 дБ / 0,33 см ²	8,3 дБ / 9,7 см ²
3	Метод градиентного спуска	16,6 дБ / 0,33 см ²	15,9 дБ / 0,062 см ²	9,1 дБ / 0,9 см ²

Литература

1. Lashkari A.H., Parhizkar B. and others. WiFi-Based Indoor Positioning System // Computer and Network Technology (ICCNT). 2010 Second International Conference. 23–25 April 2010. pp. 76–78.
2. Frost C. and others. Bluetooth Indoor Positioning System Using Fingerprinting. In: J. Del Ser, E. Axel Jorswieck, J. Miguez, M. Matinmikko, D.P. Palomar, S. Salcedo-Sanz, S. Gil-Lopez (eds.) Mobile Lightweight Wireless Systems. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, Springer Berlin Heidelberg, 2012. vol. 81. pp. 136–150.
3. Щекотов М.С., Кашевник А.М. Сравнительный анализ систем позиционирования в помещениях, основанных на технологиях связи, поддерживаемых смартфонами // Тр. СПИИРАН. 2012. Вып. 23. С. 459–471.
4. Абдрахманова А.М., Намиот Д.Е. Использование двумерных штрихкодов для создания системы позиционирования и навигации в помещении // Прикладная информатика. 2013. № 1. С. 31–39.
5. Кий К.И., Смирнов А.М. Автономная навигация в помещениях робота «Амур» по цветовым меткам // Электронный научно-технический журнал «Техническое зрение». 2013. Вып. 2. С. 30–36.
6. Бабаян П.В., Алпатов Б.А. Выделение движущихся объектов в условиях геометрических искажений изображений // Цифровая обработка сигналов. 2004. № 4. С. 9–14.
7. Белобрюхов М.С., Романенко А.В. Система машинного зрения для киберфутбола // Доклады ТУСУРа. Декабрь. 2011. № 2 (24). часть 2. С. 200–203.
8. Дворкович В.П., Дворкович А.В. Метрологическое обеспечение видеоинформационных систем. – М.: Техносфера, 2015. – 784 с.
9. Лебедев И.М., Тюкин А.Л., Приоров А.Л. Разработка и исследование системы навигации внутри помещений для мобильного робота, с возможностью детектирования препятствий // Информационно-измерительные и управляющие системы. 2015. Т. 13, № 1. С. 53–61.
10. Джакония В.Е. Телевидение. – М.: Горячая линия – Телеком, 2002. – 640 с.
11. Дворкович В.П., Дворкович А.В. Цифровые видеоинформационные системы (теория и практика). – М.: Техносфера, 2012. – 1009 с.
12. Костильов В.П., Слусар Т.В., Суший А.В., Черненко В.В. Про покращення чутливості кремнієвих фотосенсорів // Прикладна радіоелектроніка: наук.- техн. журнал. 2012. Т. 11, № 3. С. 440–444.
13. Tyukin A., Lebedev I., Priorov A. The development and research of the indoor navigation system for a mobile robot with the possibility of obstacle detection // Open Innovations Association (FRUCT16). 2014. 16th Conference of, pp. 115–122. 27–31 Oct. 2014.
14. Тюкин А.Л., Лебедев И.М., Приоров А.Л. Разработка и оценка качества работы алгоритма цифровой обработки телевизионных изображений для задач позиционирования в замкнутом пространстве // Нелинейный мир. 2014. Т. 12, № 12. С. 26–30.
15. Priorov A., Tumanov K., Volokhov V. Efficient Denoising Algorithms for Intelligent Recognition Systems. In: Favorskaya M., Jain L.C. (eds.) Computer Vision in Control Systems 2, Intelligent Systems Reference Library, Vol. 75, Springer International Publishing, Switzerland, 2015. pp. 251–276.
16. Шапиро Л., Стокман Дж. Компьютерное зрение. – М.: БИНОМ. Лаборатория знаний, 2006. 752 с.
17. Бабаян П.В., Алпатов Б.А. Методы обработки и анализа изображений в бортовых системах обнаружения и сопровождения объектов // Цифровая обработка сигналов. 2006. № 2. С. 45–51.
18. Гонсалес Р., Вудс Р. Цифровая обработка изображений. – М.: Техносфера, 2005. 1104 с.
19. Апальков И.В. Усовершенствование алгоритмов удаления шума из изображений на основе модифицированных критериев оценки качества: автореф. дис. к-та тех. наук: 05.12.04 / И.В. Апальков. М., 2008. 24 с.
20. Приоров А.Л., Куйкин Д.К., Хрящев В.В. Детектирование и фильтрация импульсного шума со случайными значениями импульсов // Цифровая обработка сигналов. 2010. № 1. С. 18–22.
21. Волохов В.А. Подавление гауссовского шума в изображениях на основе анализа главных компонент и нелокальной обработки: автореф. дис. к-та тех. наук: 05.12.04 / В.А. Волохов. Владимир, 2012. 19 с.
22. Тюкин А.Л. О скорости цифровой обработки телевизионного изображения для алгоритма позиционирования в помещении // Докл. междунар. конф. «Радиоэлектронные устройства и системы для инфокоммуникационных технологий – РЭУС-2015». 2015. С. 300–304.

НОВЫЕ КНИГИ

Авиационные системы радиовидения / Под ред. Г.С. Кондратенкова. – М.: Изд-во «Радиотехника», 2015г. – 648 с.: ил.

Дано систематическое изложение основ теории авиационных систем радиовидения (амплитудные, фазовые, скоростные и поляризационные портреты, многопозиционные системы, обратное синтезирование). Приводятся примеры реализации и эффективности различных методов радиовидения. Книга написана с использованием открытых опубликованных трудов научной школы «Радиовидение» ВВИА им. проф. Н.Е. Жуковского.

Для научных работников и инженеров, а также преподавателей, аспирантов и студентов высших учебных заведений радиотехнического профиля.

Лавров А.А. Авиационные обзорные радиолокаторы. Методы и алгоритмы пространственно-временной обработки сигналов. Монография. – М.: Изд-во «Радиотехника», 2015 г. – 224 с.: ил.

Рассмотрены принципы построения, методы и алгоритмы обработки сигналов, использующих пространственно-временную обработку сигналов. Показаны их возможности при обзоре земной и морской поверхности в режимах синтезирования апертуры антенны и формирования скоростного портрета. Приведены алгоритмы формирования изображения, технические требования к аппаратуре и характеристики создаваемых портретов – пространственная разрешающая способность и точность измерения скорости.

УДК 004.932.2, 004.932.4, 004.925.5

ОЦЕНКА ЭФФЕКТИВНОСТИ СПОСОБОВ ЦВЕТОВОЙ КОРРЕКЦИИ ДЛЯ ПАНОРАМНЫХ ИЗОБРАЖЕНИЙ С МАЛОРАЗМЕРНЫМИ ОБЪЕКТАМИ

Сильвестрова О.В., инженер ОАО «Центр наукоёмких технологий», e-mail: silvol@mail.ru.

COLOR CORRECTION METHODS EFFICIENCY EVALUATION FOR PANORAMIC IMAGES WITH SMALL-SIZE OBJECTS

Silvestrova O.V.

Color correction methods for panoramic images with small-size objects analysis implemented. Qualitative and quantitative efficiency assessment of color correction methods for tasks with small-size objects performed. Local color correction methods advantage shown.

Keywords: color correction, structural similarity, color similarity, local color correction, global color correction, parametric methods.

Ключевые слова: цветовая коррекция, структурное подобие, цветовое подобие, локальные методы цветовой коррекции, глобальные методы цветовой коррекции.

Введение

Одной из наиболее распространенных проблем в обработке панорамных изображений является различие в уровне интенсивности и цветности сшиваемых изображений, которое возникает вследствие разных уровней экспозиции (освещенности), разницы в углах зрения и др. Чаще всего эта проблема решается применением метода смешивания или метода компенсации в зоне перекрытия [1, 2]. Однако при задачах обнаружения, сопровождения малоразмерных объектов применение данных методов может привести к потерям объектов [3].

Методы цветовой коррекции

Предложенные в литературе методы цветовой коррекции можно разделить на параметрические подходы с привлечением моделей [4] и непараметрические подходы без использования моделей [5]. Обычно, параметрические подходы работают лучше их непараметрического оппонента [6]. Параметрические подходы основаны на работе [7], где для цветовой коррекции данного изображения (обозначенного как целевое изображение) была предложена простая статистическая структура, использующая цвета другого изображения (обозначенного как модель).

В целевом изображении I_t перенесенный цвет на пиксель C_t в цветовом пространстве lab выражается как:

$$g(C_t) = \mu_s + \frac{\sigma_s}{\sigma_t} (C_t - \mu_t), \quad (1)$$

где μ_s и μ_t – математические ожидания гауссовых распределений в цветовом пространстве lab соответственно исходного и целевого изображений, σ_s и σ_t соответствующие стандартные отклонения. Между осями в пространстве lab существует небольшая корреляция, которая позволяет применять различные операции в различных цветовых каналах с некоторой уверенностью, что не

Проведен анализ методов цветовой коррекции панорамных изображений с малоразмерными объектами. Выполнена сравнительная качественная и количественная оценка эффективности способов цветовой коррекции для работы с малоразмерными объектами. Показано преимущество локальных методов цветовой коррекции.

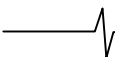
возникнет нежелательных межканальных артефактов. Кроме того, это цветовое пространство логарифмическое, что в первом приближении означает то, что равномерные изменения интенсивности в канале будут равным образом различимыми.

Другие более общие подходы глобального моделирования включают работы Tian с соавторами [8], которые используют для вычисления матрицы преобразования составление карты гистограммы в перекрывающейся области, и работу Zhang с соавторами [9], применяющими для вычисления матрицы преобразования карты «главных» областей, где главными областями обозначены области максимумов в гистограмме оттенков.

Однако очевидно, что вследствие различных отражающих свойств поверхностей и неравномерной освещенности глобальная статистика по всему изображению даст только грубое приближение всех функций цветового переноса. Глобальный перенос цвета не имеет адекватного пространственного рассмотрения, поэтому не может избежать двух следующих проблем. Первой проблемой является то, что если исходное или целевое изображение содержит разные цветовые области, то глобальный перенос не может отличить различную статистику и смешивает эти области. Второй проблемой является то, что если цвета двух изображений сильно отличаются, то в цветовом пространстве lab цветовые каналы становятся преувеличенными, что в результате приводит к неестественным и насыщенным (цветам).

Tai и др. [10] предложили локальную схему цветового переноса, основанную на вероятностной сегментации изображения и картированию областей с использованием моделей Гауссовского смешивания.

Продолжением исследований в данном направлении стала работа Oliveira [11], в которой был предложен но-



вый способ параметрической локальной цветовой коррекции. Первоначально вычисляются несколько функций передачи цвета из выходных данных алгоритма сдвига среднего математического ожидания цветовой сегментации. Затем вычисляются карты влияния цветов. Наконец, вклад каждой функции передачи цвета смешиваются с использованием весовых коэффициентов, полученных из карт влияния цветов.

Подобным образом Kim и Pollefeys [12] предложили схему максимизации подобия для устойчивого вычисления функции переноса яркости (ФПЯ) из двумерной объединенной гистограммы интенсивности двух перекрывающихся изображений. На практике метод применен к трем цветовым каналам, каждый канал имеет собственную ФПЯ. Для поиска устойчивого вычисления при условии обеспечения монотонности было применено динамическое программирование. Вычисленная ФПЯ далее использовалась для вычисления и устранения разницы в экспозиции и эффекта виньетирования в изображениях.

Таким образом, хотя в последнее десятилетие были предложены и широко используются различные технологии цветовой коррекции, сравнительной оценки эффек-

тивности этих способов для работы с малоразмерными объектами, насколько нам известно, не проводилось. Большинство авторов либо демонстрировали свои системы на нескольких ими самими выбранных примерах, либо сравнивали с очень простыми основными подходами. Для сравнительной оценки нами были выбраны как глобальные методы (метод №1 [7]), работающие в разных цветовых пространствах (метод №2 [13]), так и локальные методы (метод №3 [11], метод №4 [10]) использующие разные вероятностные характеристики также нами был исследован метод [14], основанный на тензорном голосовании, однако из-за низкой скорости вычислений он был исключен из дальнейшего рассмотрения. Для оценки результатов были созданы наборы изображений как синтетических с отличающимися яркостными характеристиками по границе +6, +8, +10, +12, так и естественные. Также в наборах изображения отличались по сюжету с различным процентным содержанием площади неба 75 %, 50 %, 30 % 25 %. В наборах изображений присутствовали малоразмерные объекты (самолет) с разным размером по отношению к площади всего кадра. Пример пары изображений с процентным отношением неба 25 % представлен на рис. 1.



Рис. 1. Пример пары изображений с процентным отношением неба 25 %

Методика сравнительной оценки

Недавно предложенная теория оценки качества изображения [15] утверждает, что с точки зрения восприятия качество результирующего изображения с измененным цветом должно показывать как цветовую, так и структурную согласованность, поскольку цветовая коррекция может не только изменять цвет результирующего изображения, то также и его структуру. Базируясь на этой теории, рассмотрим критерии оценки качества цветовой коррекции исходного изображения s в первичное панорамное изображение t , что приводит к скорректированному изображению r . Предложенный критерий оценки включает два компонента: цветовое подобие $CS(r,s)$ между исходным изображением s и перенесенным изображением r , и структурное подобие $SS(r,t)$ между результирующим изображением t и перенесенным изображением r . Цветовое подобие $CS(r,s)$ определяется как:

$$CS(r,s) = PSNR(r_1,s_1), \quad (2)$$

где $PSNR = 20 \lg(L/RMS)$ пиковое отношение сигнал/шум, L – наибольшая возможная величина в динамическом диапазоне изображения, а RMS среднеквадратичная разница между двумя изображениями, r_1 и s_1 – перекрывающиеся области r и s соответственно. Структурное подобие $SS(r,t)$ определяется как:

$$SS(r,t) = SSIM(r,t), \quad (3)$$

где $SSIM(r;t) = \frac{1}{N} \sum_{j=1}^N SSIM(a_j, b_j)$ есть индекс структурного подобия, N это число локальных окон для изображения, a_j и b_j то есть содержание изображения в j -м локальном окне изображений r и t соответственно. В нашей работе мы использовали установки по умолчанию, рекомендованные в [15].

Результаты приведены в табл 1.

Оценка результатов

Разброс в данных цветового подобия CS объясняется тем, что вычисление CS построено на пиковом отношении сигнал/шум $PSNR$, которое основывается на сред-

неквадратичных ошибках, это значит, что в реальном ряду изображений существует больше отличий внутри пар и между парами. По сравнению с методом №1, который работает в некоррелированном $l\alpha\beta$ цветовом пространстве, метод № 2 работает в коррелированном цветовом пространстве RGB, что ведет к ухудшениям в эффективности цветовой коррекции. Методы № 3 и № 4 используют локальную пространственную информацию для управления передачей цвета, что ведет к выигрышу в эффективности цветовой коррекции. Метод № 3 показал наибольшее быстроедействие, однако, так как он использует алгоритм сдвига математического ожидания, который недостаточно чувствителен к регистрации малоразмерных объектов, то для работы с малоразмерными объектами он не может быть рекомендован. Исследования, проводимые для рядов изображений с разными по размерам, по отношению к площади всего кадра объектам, показали, что изменение размеров в пределах 0,25 %-0,1 % дает схожие результаты, поэтому в табл. 1. представлены результаты для фиксированного размера объекта (0,25 % от площади всего кадра). Наряду с перечисленными методами были проведены исследования метода [12], которые показали, что практически для всех сцен он демонстрирует неустойчивую работу (интересующий нас объект пропал), поэтому в таблице результатов метод [12] не представлен.

Заключение

Проведен качественный и количественный анализ методов цветовой коррекции панорамных изображений с малоразмерными объектами. Показано, что в целом локальные методы цветовой коррекции [10, 11] дают лучшие результаты. Метод № 3 [11] несмотря на большее быстроедействие не может быть рекомендован для сюжетов с малоразмерными объектами. Метод [14], использующий тензорный анализ не может использоваться в системах, требующих высокое быстроедействие. Метод ФПЯ [12] также работает неустойчиво для изображений с малоразмерными объектами. Таким образом, дальнейшим направлением работы должно быть создание параметрического локального метода цветовой коррекции.

Таблица 1. Сравнительная оценка цветового и структурного подобия методов по [15]

№ метода	1	2	3	4
Исходное изображение Небо 75 %	CS=21,372 SS=0,892	CS=20,547 SS=0,918	CS=24,775 SS=0,967	CS=22,322 SS=0,926
Измененное изображение Яркость +6	CS=20,526 SS=0,964	CS=20,311 SS=0,856	CS=25,772 SS=0,943	CS=24,542 SS=0,998
Измененное изображение Яркость +8	CS=20,785 SS=0,974	CS=20,869 SS=0,985	CS=21,334 SS=0,871	CS=20,825 SS=0,987
Измененное изображение Яркость +10	CS=21,328 SS=0,925	CS=23,345 SS=0,982	CS=21,255 SS=0,822	CS=20,568 SS=0,879
Измененное изображение Яркость +12	CS=22,778 SS=0,882	CS=20,632 SS=0,997	CS=24,689 SS=0,891	CS=29,854 SS=0,971
Естественное изображение Небо 50 %	CS=19,113 SS=0,715	CS=19,451 SS=0,984	CS=25,638 SS=0,883	CS=22,674 SS=0,844
Естественное изображение Небо 30 %	CS=18,375 SS=0,745	CS=18,391 SS=0,741	CS=29,215 SS=0,965	CS=28,224 SS=0,979
Естественное изображение Небо 25 %	CS=19,212 SS=0,882	CS=18,217 SS=0,956	CS=26,574 SS=0,824	CS=23,513 SS=0,983

Литература

1. M. Brown and D. G. Lowe. Recognizing panoramas. In Proc. ICCV'03, volume 2, pages 1218–1225, 2003.
2. M. Brown and D. G. Lowe. Automatic panoramic image stitching using invariant features. IJCV, 74(1):59–73, 2007.
3. Обидин Г.И., Сильвестрова О.В. Способы оценки эффективности цветовой коррекции в панорамных изображениях с малоразмерными объектами // Труды 20-й Всероссийской научно-технической конференции «Современное телевидение», Москва, ФГУП МКБ «Электрон», 2012.
4. Xiang, B. Zou, and H. Li. Selective color transfer with multi-source images. Pattern Recognition Letters, 30(7): 682–689, May 2009.
5. J. Yin and J.R. Cooperstock. Color correction methods with applications to digital projection environments. Journal of WSCG, vol. 12: 1–3, 2004.
6. W. Xu and J. Mulligan. Performance evaluation of color correction approaches for automatic multi-view image and video stitching. In IEEE Int. Conference on Computer Vision and Pattern Recognition, San Francisco, USA, 2010.
7. E. Reinhard, M. Ashikhmin, B. Gooch, and P. Shirley. Color transfer between images. IEEE Computer Graphics and Applications, 21(5):. 34–41, 2001.
8. G.Y. Tian, D. Gledhill, D. Taylor, and D. Clarke. Color correction for panoramic imaging. In Proc. 6th International Conference on Information Visualization, pages 483–488, 2002.
9. M. Zhang and N.D. Georganas. Fast color correction using principal regions mapping in different color spaces. Real-Time Imaging, 10(1):23–30, 2004.
10. Y.-W. Tai, J. Jia, and C.-K. Tang. Local color transfer via probabilistic segmentation by expectation-maximization. In Proc. CVPR'05, volume 1, pages 747–754, 2005.
11. Miguel Oliveira, Angel D. Sappa Unsupervised Local Color Correction for Coarsely Registered Images IEEE conference on Computer Vision and Pattern Recognition: 201–208, 2011
12. S.J. Kim and M. Pollefeys. Robust radiometric calibration and vignetting correction. IEEE TPAMI, 30(4):562–576 2008.
13. X. Xiao and L. Ma. Color transfer in correlated color space. In Proc. 2006 ACM international conference on Virtual reality continuum and its applications, pages 305–309, 2006.
14. J. Jia and C.-K. Tang. Tensor voting for image correction by global and local intensity alignment. IEEE TPAMI, 27(1):36–50, 2005.
15. Z. Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli. Image quality assessment: From error visibility to structural similarity. IEEE Transactions on Image Processing 2004.

Уважаемые коллеги!

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

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

Научно-технический журнал «Цифровая обработка сигналов» включен в Перечень изданий, рекомендуемый ВАК РФ для публикации результатов научных исследований соискателями ученой степени доктора и кандидата технических наук в области радиотехники, связи, вычислительной техники, электроники, приборостроения, информационных технологий, информационно-измерительных и управляющих систем. По предварительным итогам за 2014 год по рейтингу Science Index базы РИНЦ (3,394) журнал «Цифровая обработка сигналов» занимает 344-ю позицию из почти 3000 представленных изданий.

Планируемые сроки издания отдельных номеров журнала:

- № 1 – апрель 2016 г. Тематический выпуск: «ЦОС в информационно-измерительных и управляющих системах»;
- № 2 – июнь 2016 г. Тематический выпуск: «Теория и методы цифровой обработки сигналов» (по материалам международной научно-технической конференции «Цифровая обработка сигналов и ее применение – DSPA' 2016»);
- № 3 – сентябрь 2016 г. Тематический выпуск: «Цифровая обработка изображений»;
- № 4 – декабрь 2016 г. Тематический выпуск: «ЦОС в радиотехнике и системах телекоммуникаций».

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

УДК 528.852.3

СТРУКТУРНОЕ ВОССТАНОВЛЕНИЕ ИЗОБРАЖЕНИЙ ОТ МНОГОМАТРИЧНЫХ СКАНЕРНЫХ ДАТЧИКОВ ДИСТАНЦИОННОГО ЗОНДИРОВАНИЯ ЗЕМЛИ

*Кузнецов А.Е., д.т.н., НИИ «Фотон» РГРТУ, e-mail: foton@rsreu.ru;
Пресняков О.А., к.т.н., НИИ «Фотон» РГРТУ, e-mail: foton@rsreu.ru;
Мятов Г.Н., к.т.н., АО «РКЦ «Прогресс».*

STITCHING OF REMOTE SENSING IMAGES FROM STAGGERED TDI CCD

Kuznetsov A.E., Presniakov O.A., Myatov G.N.

The paper is devoted to geometrical stitching of images from staggered TDI CCD. Factors affecting adjacent CCD images displacement are discussed. Stitching methods, including those which authors used for development of software for processing of imagery from «Resurs-DK», «Resurs-P» and other satellites, are reviewed and analyzed. Methods for precise stitching process organization are proposed.

Keywords:

Ключевые слова: «сшивки», ПЗС ВЗН, многоматричный датчик.

Введение

Разрешающая способность и полоса захвата современных высокоточных спутниковых систем наблюдения Земли (табл. 1) требуют формирования изображений с высокой строчной частотой и большим числом пикселей в строке (например, строка панхроматического изображения от КА «Ресурс-П» содержит 54 тысячи пикселей). Для повышения времени экспонирования и чувствительности датчика используются матрицы ПЗС, работающие в режиме временной задержки и накопления (ПЗС ВЗН, «TDI-CCD»). Имеющаяся в наличии элементная база не позволяет охватить одной ПЗС-матрицей всю полосу захвата, поэтому распространен подход, при котором для достижения требуемых характеристик систем ДЗЗ в фокальной плоскости размещают несколько ПЗС-матриц в шахматном порядке [1] (рис. 1). В иностранной литературе такое размещение могут описывать терминами «staggered», «non-collinear» [2, 3].

Статья посвящена вопросу геометрической «сшивки» изображений от многоматричных датчиков. Рассмотрены факторы, влияющие на смещение общих точек в изображениях от соседних ПЗС-матриц. Выполнен анализ методов, используемых для «сшивки», включая методы, использованные авторами при разработке программного обеспечения обработки данных от КА «Ресурс-ДК», «Ресурс-П» и других. Рассмотрены вопросы организации процесса высокоточной сшивки.

В процессе съемки каждая ПЗС-матрица формирует отдельное изображение, так называемый скан.

Концепция функционирования наземных комплексов обработки информации для всех типов КА предусматривает выдачу потребителям выходной продукции, не учитывающей специфику построения и функционирования сканирующего датчика. В нашем случае это означает, что потребителям должно выдаваться непрерывное изображение наблюдаемой территории (рис. 2), а не набор отдельных сканов. Отсюда возникает задача высокоточного структурного восстановления информации или, так называемой, геометрической «сшивки».

Таблица 1. Характеристики спутников с многоматричными видеодатчиками

Спутник	Страна	Год запуска	Полоса захвата, км	Разрешение, м
IRS-1C	Индия	1995	70	5,8
IKONOS	США	1999	11	0,82 (ПХ), 3,2 (МС)
QuickBird 2	США	2001	18	0,61 (ПХ), 2,44 (МС)
СBERS-02В	Бразилия-Китай	2003	113	20
«Ресурс-ДК»	Россия	2006	28,3 (H=350 км)	не хуже 1
WorldView-1	США	2007	17,7	0,5
GeoEye-1	США	2008	15,3	0,41 (ПХ), 1,65 (МС)
Worldview-2	США	2009	16,4	0,46 (ПХ), 1,84 (МС)
ZY-1 02С	Китай	2011	27 (1 прибор)	2,36
«Ресурс-П» №1, №2	Россия	2013, 2014	38	не хуже 1 (ПХ)
Worldview-3	США	2014	13,1	0,31 (ПХ)

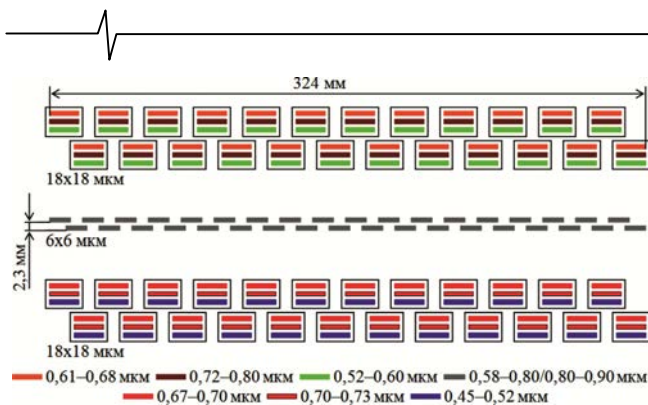


Рис. 1. Размещение ПЗС-матриц

в фокальной плоскости датчика КА «Ресурс-П»

Для последующего анализа известных технических решений проанализируем основные факторы, влияющие на характер смещений изображений от соседних ПЗС-матриц.

Структурные искажения многоскановых изображений

При съемке с использованием ПЗС ВЗН заряды в них должны переноситься параллельно движению изображения в фокальной плоскости. В противном случае будет возникать смаз изображения, а при использовании ПЗС-матриц, расположенных в шахматном порядке, возможно также снижение перекрытия между сканами вплоть до появления разрывов. Поэтому на параметры движения изображения в фокальной плоскости (а значит, на движение КА) накладываются определенные ограничения [1]. Система управления движением КА (СУД) на основе данных о его орбитальном и угловом движении, конструкции датчика, вращении Земли и т.д. должна таким образом управлять ориентацией спутника, чтобы обеспечить стабильность скорости движения и минимизировать наклон вектора скорости движения изображения относительно направления столбцов ПЗС ВЗН. Полностью влияние перечисленных

факторов скомпенсировать, как правило, невозможно или не удастся, что приводит к отклонению смещения сканов от номинального величиной несколько пикселей. Это отклонение непостоянно и плавно изменяется в пределах снимка. Влияние ряда других факторов на движение изображения и зависящее от него смещение сканов в принципе не может быть скомпенсировано СУД. К таким факторам в первую очередь относятся высокочастотное изменение углов ориентации КА и рельеф местности, приводящий к параллактическим искажениям [1, 2, 4].

Изменение углов ориентации спутника. Одна и та же точка земной поверхности фиксируется соседними ПЗС-матрицами с некоторым временным интервалом. За это время изменяются углы тангажа (α), крена (ω) и рысканья (χ) спутника, что приводит к значительным поперечным и продольным смещениям одноименных точек в соседних сканах. В первом приближении при съемке в надире изменение угла тангажа $\Delta\alpha$ (в радианах) приводит к смещениям точек на Земле вдоль направления полета $\Delta X = \Delta\alpha H$; изменение угла крена $\Delta\omega$ вызывает смещения в поперечном направлении $\Delta Y = \Delta\omega H$; изменение угла рысканья $\Delta\chi$ приводит к развороту строк на угол $\Delta\chi$ и смещению вдоль направления полета на разную величину в пределах снимка.

Рельеф местности. Рассмотрим, как влияет отклонение Δh высоты точек от некоторой средней геодезической высоты в пределах снимка \bar{h} на смещения соседних сканов в этих точках (рис. 3). Если точка T на поверхности Земли имеет высоту \bar{h} , она регистрируется соседними ПЗС-матрицами в положениях на орбите 1 и 2. При высоте точки $\bar{h} + \Delta h$, ее смещение в соседних сканах изменится на величину, эквивалентную расстоянию ε на поверхности Земли.

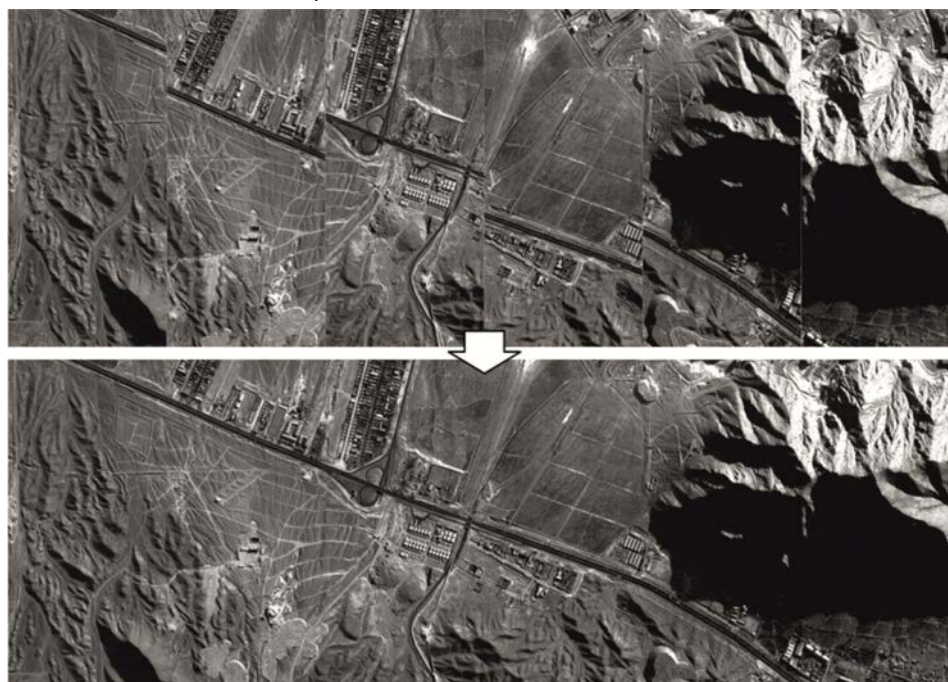


Рис. 2. Многоскановое и структурно восстановленное изображения

В первом приближении, рассматривая надирную съемку и пренебрегая кривизной поверхности Земли и участка орбиты в районе снимка, $\varepsilon = \Delta h \frac{\ell}{f}$, где ℓ – расстояние между ПЗС-матрицами в фокальной плоскости, f – фокусное расстояние оптической системы. Часто удобнее использовать оценку смещения в пикселях $k = \frac{\Delta h \ell}{H r}$, где H – высота полета спутника, r – размер фотозлемента ПЗС-линейки. Выразив через N номинальное смещение ПЗС-линейек в пикселях, получим $k = N \frac{\Delta h}{H}$. В табл. 2 приведены значения k для некоторых спутников при высоте полета H и перепаде высот в районе съемки 1 км.

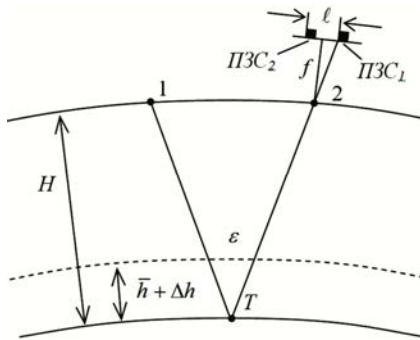


Рис. 3. Схема оценки смещения одноименных точек соседних сканов из-за рельефа местности

Таблица 2. Смещение общих точек в соседних сканах из-за рельефа местности

Датчик (КА)	N, пикс.	H, км	k, пикс.
Геотон («Ресурс-ДК»)	256	350–600	0,43–0,73
Геотон-Л1 («Ресурс-П №1»), ПХ	384	475	0,81
Геотон-Л1 («Ресурс-П №1»), МС	414	475	0,87
HR1/HR2 (ZY-1 02C)	2600	779	3,3

На рис. 4 представлены графики смещений номеров строк одноименных точек на соседних сканах, идентифицированных при помощи корреляционно-экстремаль-

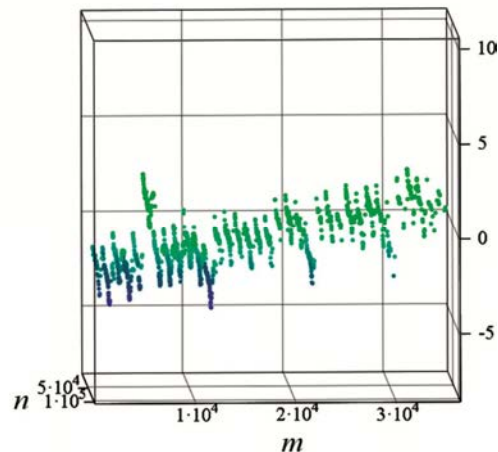
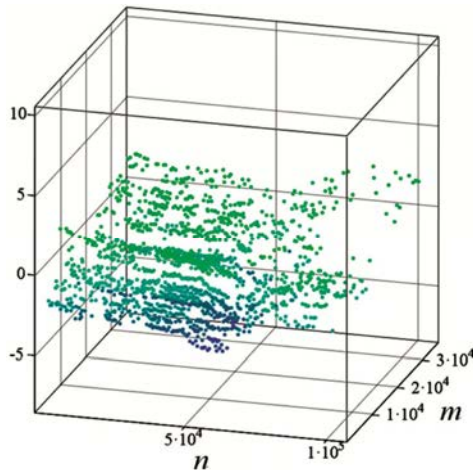


Рис. 4. Графики изменения смещения одноименных точек в пикселях в строчном направлении в пределах снимка

ной процедуры, для реального сканового изображения от КА «Ресурс-ДК». Длинные «борозды», тянущиеся в кадровом направлении и хорошо различимые на графике справа, обусловлены неточностями установки ПЗС-линейек в фокальной плоскости. Общий наклон графика определяется орбитой спутника, углами съемки и вращением Земли. Колебания углового движения приводят к волнообразным изменениям смещений в зависимости от номера строки. Рельеф местности на приведенном примере вносит незначительные локальные несистематические искажения.

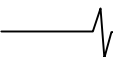
Методы «сшивки» сканов

В литературе рассмотрены различные варианты решения задачи «сшивки» изображений от многоматричных датчиков с ПЗС ВЗН [2-9]. К сожалению, в доступных источниках не для всех коммерческих спутников удалось найти описание используемых алгоритмов «сшивки».

«Сшивка» плоскопараллельным сдвигом. Из анализа факторов, влияющих на перекрытие соседних сканов, понятно, что субпиксельной точности «сшивки» этим методом можно достичь только при условии максимально близко расположенных друг к другу в кадровом направлении ПЗС-матриц, поэтому для большинства эксплуатируемых сегодня датчиков с шахматным расположением ПЗС ВЗН его применение нецелесообразно.

Методы «сшивки» на основе определения перекрытий по видеоданным. В литературе [2] для обозначения таких методов также встречается термин «ориентированные на пространство изображения» («image-space-oriented»).

В работе [5] описано получение функции смещения соседних сканов на основе оригинального метода функционализации параметров изображений. К достоинствам метода можно отнести возможность достижения точности совмещения изображений в области шва лучше 0,1 пикселя. Для работы алгоритма не требуется высокоточных измерений положения и ориентации КА. Авторы отмечают высокое быстродействие алгоритма, позволяющее использовать его на борту. К сожалению, в работе не приводится алгоритм собственно «сшивки» на основе полученной функции.



В работе [3] рассмотрен метод «сшивки» путем заполнения пробела между видеоданными нечетных матриц информацией от четных, используемой в качестве «заплатки». Рассматривая процесс «сшивки» изображений от трех ПЗС-матриц, авторы вводят понятие виртуальной ПЗС-матрицы ВПЗС₂, закрывающей пробел между ПЗС₁ и ПЗС₃ и расположенной на одной линии с ними (рис. 5). Трансформирование координат между ПЗС₂ и ВПЗС₂ выполняется по кусочно-полиномиальной модели, параметры которой определяются на основе связующих точек. Идентификацию связующих точек предлагается выполнять по алгоритму SIFT с последующей фильтрацией.

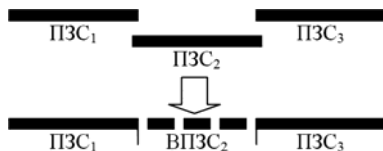


Рис. 5. Размещение виртуальной ПЗС-матрицы в методе заполнения пробелов

Область применения методов этой группы имеет ограничения, которые накладывает сюжет снимков. Так, могут возникнуть сложности с обработкой не подвергающихся идентификации сюжетов: содержащих преимущественно воду, снежные поля и т.п.

Другим общим недостатком методов «сшивки» на основе определения перекрытий по стыкам является то, что положение точек, удаленных от краев сканов, не будет зависеть от рельефа местности в этих точках. Использование таких моделей в случае, когда расстояние между ПЗС-матрицами в кадровом направлении составляет сотни пикселей, может привести к ошибкам геопривязки «сшитого» изображения. Величина ошибок будет соответствовать параллактическому смещению точек в соседних сканах (табл. 2).

Регрессионный метод «сшивки», впервые предложенный в 2004 г. [6], позднее доработанный [7] и подробно рассмотренный в литературе [4], предполагает определение на основе общих точек параметров модели, адекватно описывающей трансформирование сканов в пределах всего структурно восстановленного изображения. За счет избыточности точек по отношению к неизвестным параметрам функции достигается высокая надежность метода; он нечувствителен к локальным ошибкам идентификации и невозможности идентифицировать общие точки в локальных областях. «Сшитое» изображение представляет собой изображение от виртуальной (несуществующей) ПЗС-линейки, размещенной в фокальной плоскости датчика и перекрывающей совокупное поле зрения всех ПЗС-матриц, что позволяет получить преимущества, которые дает использование строгой модели.

Для трансформирования устанавливается следующее соответствие между координатами (m_k, n_k) , $k = \overline{1, K}$, точек сканов и координатами (m, n) структурно восстановленного изображения. При переносе пикселей нечетных сканов используются только установочные параметры. Для четных сканов ($k = 2, 4, \dots$) трансформирование выполняется по формуле (рис. 6):

$$m_k(m, n) = U_{mk}(m) + P_m[m, t(n)] + A_m[t(n)],$$

$$n_k(m, n) = t_k^{-1} [U_{nk}(m) + P_n[m, t(n)] + A_n[t(n)]]. \quad (1)$$

Здесь:

– $U_{mk}(m)$, $U_{nk}(m)$ – функции компенсации ошибок совмещения сканов из-за неточности установки ПЗС-матриц в фокальной плоскости;

$$P_m(m, t) = a_0 + a_1 t + a_2 m + a_3 t^2 + a_4 m^2,$$

$P_n(m, t) = b_0 + b_1 t + b_2 m + b_3 t^2 + b_4 m^2$ – полиномы, компенсирующие отклонение видеоданных отдельных ПЗС-линеек от номинального положения из-за орбитального движения КА, его ориентации и вращения Земли;

– $A_m(t)$, $A_n(t)$ – кусочно-линейные функции, учитывающие высокочастотные изменения углов крена и тангажа КА;

– $t(n)$, $t^{-1}(t)$ – прямая и обратная функции, устанавливающие связь между номером строки и временем ее съемки, необходимость введения которых обусловлена возможностью смены частоты сканирования в процессе съемки.

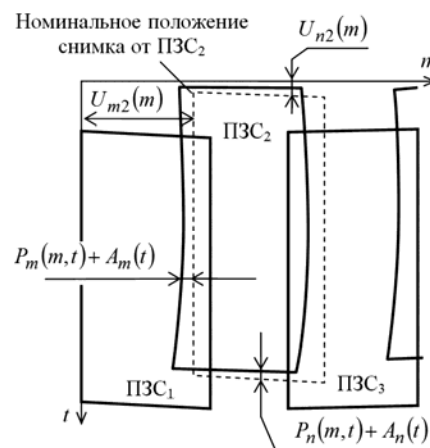


Рис. 6. Параметры, определяющие взаимное положение сканов в регрессионном методе «сшивки»

Функции $U_{mk}(m)$, $U_{nk}(m)$ определяются в результате калибровки относительного положения ПЗС-матриц. Параметры функций P_m , P_n , A_m , A_n определяются по общим точкам, для поиска которых используется корреляционно-экстремальная идентификация.

Метод успешно применяется при обработке снимков с перепадами высот до 2000 м от КА «Ресурс-ДК», для которого точность определения элементов внутреннего ориентирования съемочной аппаратуры приводит к несогласованию геодезических координат общих точек на соседних сканах до 3–5 пикселей.

Регрессионный метод, как и методы на основе определения перекрытий, снижает точность геопривязки на величину параллактических искажений между соседними сканами. Также возникают сложности с обработкой сюжетов, включающих плохо идентифицируемые области, но при условии, что сцена состоит преимущественно из таких областей. Следует отметить невозможность применения или снижение эффективности рассмотренного метода при небольшом количестве ПЗС-матриц в составе датчика (две или три).

Методы сшивки на основе точной геодезической привязки. Аппаратура современных систем ДЗЗ позволяет после ее геометрической калибровки достичь точности геопривязки, при которой рассогласование геодезических координат общих точек в соседних сканах не превышает 0,3 пикселя. В этом случае представляется возможным выполнить точное структурное восстановление в автоматическом режиме. Использование строгой геометрической модели датчика ряд авторов отмечают как необходимое условие для обеспечения высокой измерительной точности изображений [2, 8].

Один из вариантов высокоточной «сшивки» рассмотрен в работе [8]. Автором предлагается применить процедуру маршрутной фототриангуляции с использованием опорных и связующих точек, в результате которой минимизируется разность геодезических координат в общих точках сканов. Связующие точки предлагается определять автоматически с помощью цифрового коррелятора. Подход безусловно правильный, но, к сожалению, автор не уделил внимания решению собственно задачи «сшивки», т.е., получения структурно восстановленного изображения. На основе точной геопривязки можно выполнить трансформирование или ортотрансформирование сканов и сформировать «сшитое» изображение в системе координат картпроекции, но это не удовлетворяет в полной мере требованиям сложившейся международной практики, в соответствии с которыми потребителю могут также выдаваться исходные снимки с большой полосой обзора, геопривязка которых выполняется по RPC или строгой модели. Можно также отметить потенциальную чувствительность метода к ошибкам из-за идентификации связующих точек на облачности.

В работе [2] «сшитое» изображение помещается в плоскость проекции вдоль траектории съемки, которая задается по положению первого и последнего пикселей первого столбца первого скана. Сканы трансформируются в «сшитое» изображение по данным геопривязки на средней высоте местности. Структурно восстановленное изображение сопровождается коэффициентами дробных полиномов. Метод используется в штатной системе обработки изображений от спутника ZY-1 02C для «сшивки» снимков с перепадом высот до 300 м. При более сильных перепадах авторы рекомендуют использовать дополнительную геометрическую коррекцию на основе методов, ориентированных на пространство изображения. Отметим, что в предложенном варианте решения можно было бы использовать ортотрансформирование сканов, и даже с использованием достаточно грубой ЦМР происходила бы точная «сшивка» снимков горных районов. Однако использование грубой ЦМР резко отрицательно сказалось бы на измерительной точности выходной продукции.

Идея метода «сшивки» на основе точной геопривязки, обобщенного на двумерный и трехмерный (с использованием ЦМР) случай, предложена авторами в работах [4, 9]. «Сшитое» изображение предлагается считать сформированным виртуальной ПЗС-линейкой, перекрывающей совокупную полосу обзора ПЗС-матриц датчика. При такой постановке оказывается возможным использовать достаточно грубую ЦМР только для устранения межска-

новых параллактических искажений, когда ошибки высоты 300 – 1000 м в зависимости от датчика являются приемлемыми. «Сшитое» изображение может распространяться потребителям в сопровождении коэффициентов RPC либо параметров строгой модели, при этом в дальнейшем потребители могут применять свои высокоточные ЦМР для обработки интересующих их участков либо выполнять ориентирование по высокоточной опорной информации. Рассмотрим алгоритм реализации этого метода применительно к обработке с использованием ЦМР, поскольку ЦМР с грубым разрешением доступна для всей территории Земли. Кроме того, снижение ошибки «сшивки» даже на 0,1–0,2 пикселя также целесообразно, поскольку рельефная составляющая является лишь одной из частей общей ошибки. Алгоритм апробирован при обработке снимков от российского КА «Ресурс-П».

Будем считать, что исходные сканы приведены к единой частоте сканирования и номер строки n_k линейно связан с временем t . Обозначим вектор координат точки на k -м скане, дополненных высотой над поверхностью эллипсоида, как $\mathbf{r}_k = (m_k, n_k, h)$, $k = \overline{1, K}$. Вектор-функцию геопривязки, определяющую в указанной точке для заданной высоты геодезические координаты $\mathbf{g} = (B, L, h)$ (с тождественным преобразованием высоты), обозначим через \mathbf{G}_k ,

$$\mathbf{g} = \mathbf{G}_k(\mathbf{r}_k).$$

Аналогично определим вектор-функцию геопривязки для структурно восстановленного изображения как $\mathbf{g} = \mathbf{G}(\mathbf{r})$, $\mathbf{r} = (m, n, h)$.

Тогда прямое координатное соответствие исходных сканов и структурно восстановленного изображения можно представить в виде функции $\mathbf{T}_k(\mathbf{r}_k)$,

$$\mathbf{r} = \mathbf{T}_k(\mathbf{r}_k) = \mathbf{G}^{-1}[\mathbf{G}_k(\mathbf{r}_k)],$$

где $\mathbf{G}^{-1}(\mathbf{g})$ – вектор-функция, обратная функции $\mathbf{G}(\mathbf{r})$. В аналитическом виде такое обращение часто невозможно выполнить. Поэтому можно прибегнуть к численному обращению методом Ньютона.

Формирование пикселей структурно восстановленного изображения с использованием $\mathbf{T}_k(\mathbf{r}_k)$ требует сложной процедуры интерполяции в нерегулярной решетке [10], поэтому на практике обработка строится с использованием обратных функций геометрического соответствия. Их можно получить двумя способами. Во-первых, воспользоваться преобразованием, аналогичным \mathbf{T}_k , т.е. $\mathbf{r}_k = \mathbf{G}_k^{-1}[\mathbf{G}(\mathbf{r})]$. Для организации скоростной обработки также необходимо выполнить кусочно-линейную аппроксимацию полученной зависимости. При такой обработке могут возникнуть сложности при численном обращении \mathbf{G}_k , связанные с необходимостью расширения измерений за пределы скана и нахождением области определения скана в плоскости структурно восстановленного изображения. К численному обращению функции \mathbf{G} это относится в меньшей степени. Поэтому для организации скоростной обработки по обратному закону геометрического соответствия выполним кусочно-линейную аппроксимацию функции

$T_k(\mathbf{r}_k)$ функцией $\hat{T}_k(\mathbf{r}_k)$, после чего функцию $\hat{T}_k^{-1}(\mathbf{r})$, обратную $\hat{T}_k(\mathbf{r}_k)$, аппроксимируем, в свою очередь, кусочно-линейной вектор-функцией $P_k(\mathbf{r})$, $\mathbf{r}_k = P_k(\mathbf{r})$.

Далее для учета рельефа местности, получим для каждой точки структурно восстановленного изображения компоненту h вектора \mathbf{r} и выполним трехмерную кусочно-линейную аппроксимацию $T_k(\mathbf{r}_k)$ и $\hat{T}_k^{-1}(\mathbf{r})$.

Для определения высоты в каждой точке структурно восстановленного изображения строится равномерная решетка, шаг которой определяется размером пикселя ЦМР. В узлах решетки находится пересечение сканирующего луча с поверхностью Земли при помощи функции $\mathbf{g} = \mathbf{G}_{ЦМР}(m, n)$. Каждая такая операция является сложной в вычислительном отношении, но требования к точности ЦМР невысоки и при использовании ЦМР низкого разрешения операция трансформирования лишь незначительно увеличивает общее время обработки. Восстановленную при помощи билинейной интерполяции непрерывную функцию высоты обозначим $\tilde{H}(m, n)$. При заполнении высот в узлах решетки также определяются минимальная и максимальная высоты, которые необходимы для аппроксимации и обращения $T_k(\mathbf{r}_k)$.

При трехмерной кусочно-линейной аппроксимации функции $T_k(\mathbf{r}_k)$ в пространстве (m_k, n_k, h) формируется равномерная трехмерная прямоугольная сетка. Ячейки последней – прямоугольные параллелепипеды – разбиваются каждая на 6 треугольных пирамид (рис. 7). Размеры ячеек решетки определяются исходя из требуемой точности аппроксимации [11] с учетом того, что обработка предполагает еще последующую аппроксимацию $\hat{T}_k^{-1}(\mathbf{r})$. В узлах сетки по $T_k(\mathbf{r}_k)$ рассчитываются координаты в пространстве (m, n, h) . По этим узлам $T_k(\mathbf{r}_k)$ в пределах пирамид аппроксимируется трех-

мерными линейными функциями, и находятся обратные функции вида:

$$m_k = a_{ijc0} + a_{ijc1}m_k + a_{ijc2}n_k + a_{ijc3}h,$$

$$n_k = b_{ijc0} + b_{ijc1}m_k + b_{ijc2}n_k + b_{ijc3}h.$$

Эти функции используются для заполнения узлов трехмерной решетки в пространстве (m, n, h) при кусочно-линейной аппроксимации $\hat{T}_k^{-1}(\mathbf{r})$. Учитывая, что высота h является общей для (m_k, n_k, h) и (m, n, h) , можно существенно упростить задачу, выбрав шаг сеток по высоте одинаковым, минимальным из рассчитанных для аппроксимации $T_k(\mathbf{r}_k)$ и $\hat{T}_k^{-1}(\mathbf{r})$. Тогда узлы решетки для аппроксимации $T_k(\mathbf{r}_k)$, пересчитанные в (m, n, h) , будут лежать в тех же плоскостях, что и узлы решетки для аппроксимации $\hat{T}_k^{-1}(\mathbf{r})$, и расчет узлов трехмерной сетки сведется к расчету узлов двумерных треугольных сеток для каждого уровня высоты.

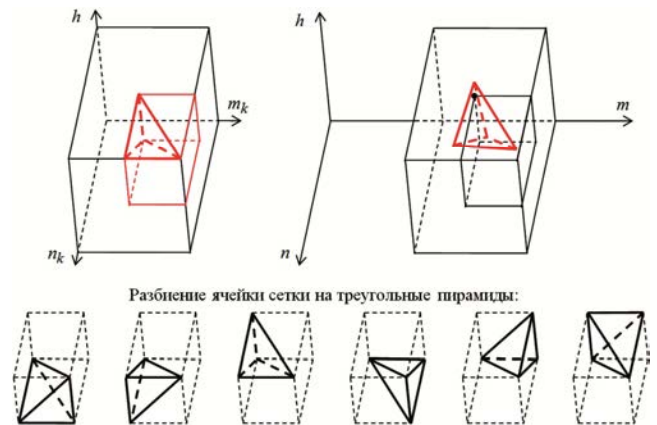
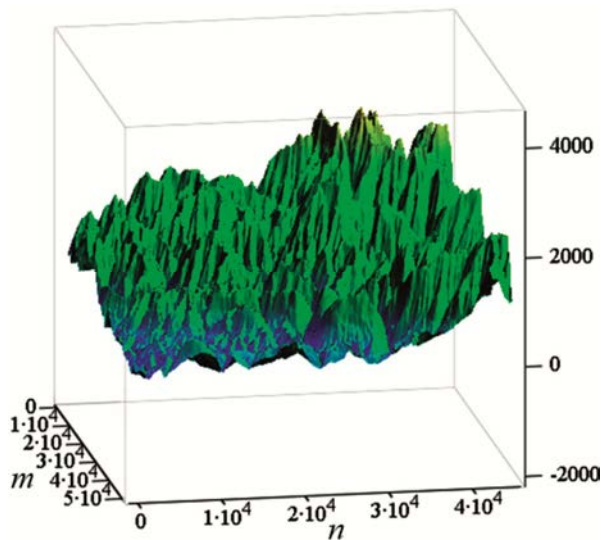
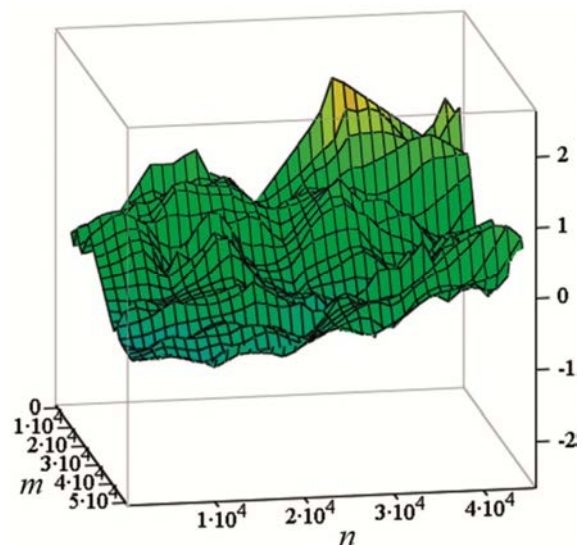


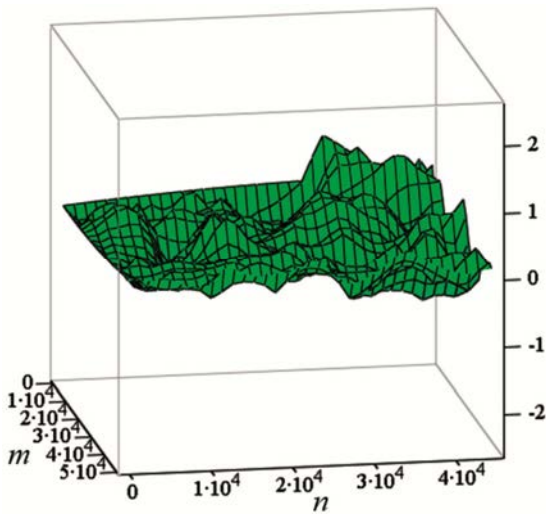
Рис. 7. Схема трехмерной аппроксимации функций геометрического соответствия сканов и структурно восстановленного изображения



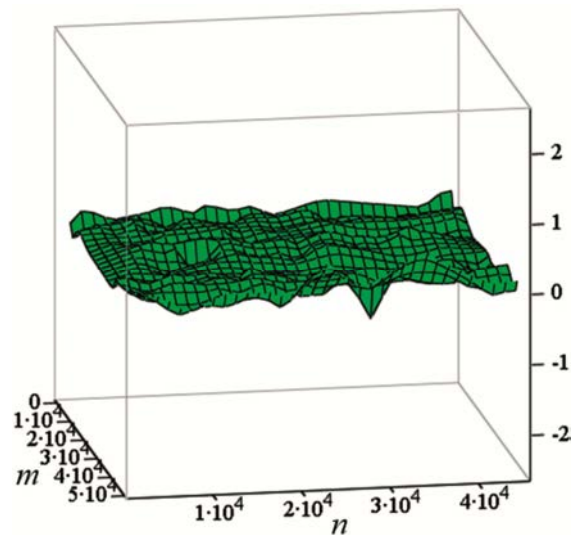
а) высота в плоскости структурно восстановленного изображения



б) ошибка совмещения без использования ЦМР



в) ошибка совмещения с использованием ЦМР с разрешением 30 угл. с



г) ошибка совмещения с использованием ЦМР с разрешением 3 угл. с

Рис. 8. Графики ошибок совмещения сканов в кадровом направлении

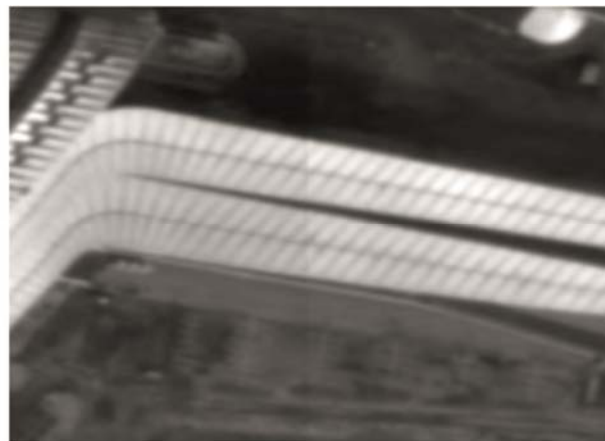
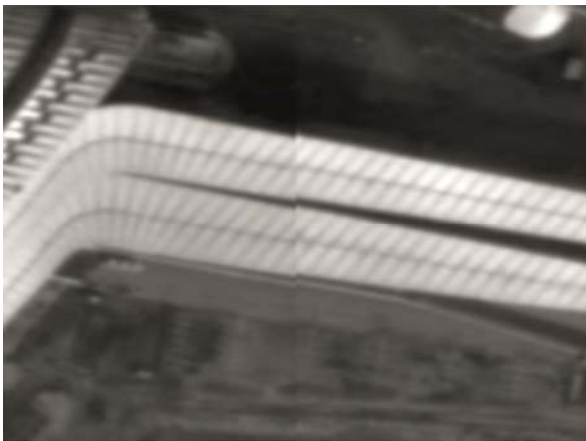


Рис. 9. Фрагменты структурно восстановленных изображений: слева – без ЦМР; справа – с ЦМР

При заполнении строки «сшитого» изображения для каждой точки (m, n) , дополненной высотой $\tilde{H}(m, n)$, определяется пирамида, в которую она попадает, и координаты исходного скана r_k .

На рис. 8 приведены графики ошибки совмещения сканов изображения от КА «Ресурс-П» в контрольных точках, иллюстрирующие эффект от привлечения ЦМР для структурного восстановления. Контрольные точки идентифицированы на соседних сканах при помощи корреляционно-экстремального анализа. Перепад высот в пределах снимка составляет 3192 м, среднеквадратическая ошибка совмещения в кадровом направлении равняется 0,57 пикселя без использования ЦМР и 0,15 пикселя при использовании ЦМР с разрешением 3 угл. с. Фрагмент этого снимка (вокзал «Олимпийский парк», г. Сочи), структурно восстановленного с использованием ЦМР и без использования ЦМР, приведен на рис. 9. Как видим, привлечение ЦМР дает более качественный результат структурного восстановления.

Заключение

Рассмотренный в работе алгоритм практически реализован в составе программного комплекса создания

стандартных продуктов по данным аппаратуры «Сангур» КА «Ресурс-П» №1 и №2. С 2013 г. этот программный комплекс успешно эксплуатируется в Научном центре оперативного мониторинга Земли, г. Москва. При этом отмечена одна важная особенность. Измерительная система спутника в ряде случаев не позволяет получить высокоточные данные об угловом движении КА. Это приводит к снижению точности строгой модели съемки и, как следствие, погрешности «сшивки» до 1–1,5 пикселя. Для выявления подобных ситуаций и организации высокоточной «сшивки» предлагается следующая технологическая схема (рис. 10).

В начале из набора сканов по строгой модели с использованием ЦМР формируется «сшитое» изображение. Для обеспечения высокой скорости обработки изображение формируется виртуально. Затем выполняется оценка точности «сшивки» на основе вычисления рассогласования координат одноименных объектов в общих частях сканов. Если среднеквадратическая ошибка совмещения $< 0,5$ пикселя, то изображение создается физически в памяти компьютера. Если точность совмещения сканов неудовлетворительная, то выполняется геодезическое ориентирование маршрутов съемки. При геодезическом ориентировании используют

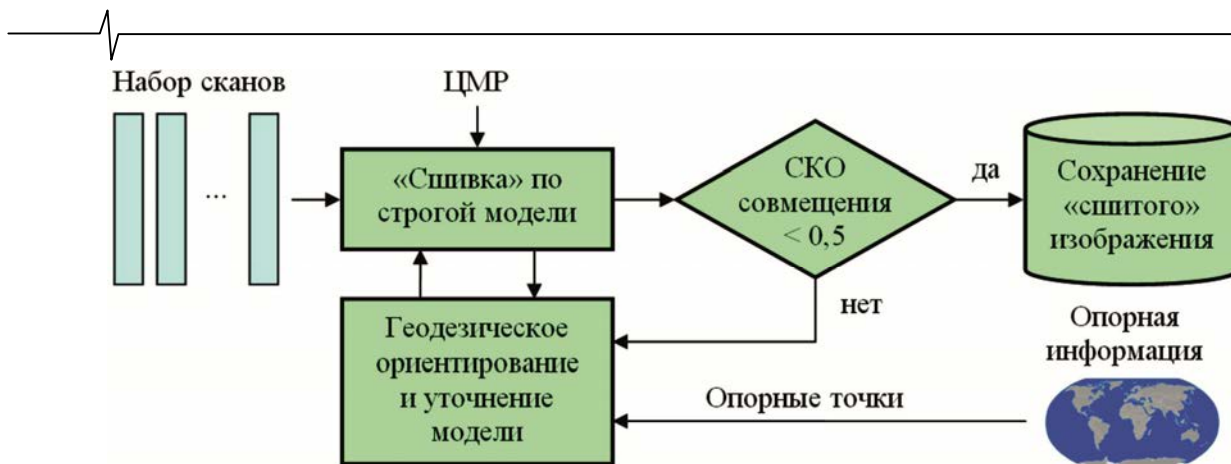


Рис. 10. Технологическая схема высокоточной «сшивки»

зуются опорные точки местности. По результатам ориентирования уточняется угловое движение КА и строгая модель съемки. Затем процесс «сшивки» повторяется, оценивается точность и синтезированное непрерывное изображение сохраняется.

Литература

1. Бакланов А.И. Системы наблюдения и мониторинга: учебное пособие / А.И. Бакланов. – М.: БИНОМ. Лаборатория знаний, 2009. 234 с.
2. Tang, X.; Hu, F.; Wang, M.; Pan, J.; Jin, S.; Lu, G. Inner FoV Stitching of Spaceborne TDI CCD Images Based on Sensor Geometry and Projection Plane in Object Space. *Remote Sens.* 2014, 6, 6386–6406.
3. Weican Meng, Shulong Zhu, Baoshan Zhu, Shaojun Bian The research of TDI-CCDs imagery stitching using information mending algorithm. *Proc. SPIE 8908, International Symposium on Photoelectronic Detection and Imaging 2013: Imaging Sensors and Applications, 89081C* (August 21, 2013); doi:10.1117/12.2033285.
4. Современные технологии обработки данных дистанционного зондирования Земли под ред. В.В. Еремеева М., Физматлит, 2015
5. Кузнецов П.К., Мартемьянов Б.В., Скирмунт В.К., Семавин В.И. Методика высокоточной сшивки изображений, получаемых при съемке обстановки многоматричным сканирующим оптико-электронным преобразователем // Вестник Самарского государственного технического университета. Серия: Технические науки. 2011. № 3 (32). С. 69–81.
6. Гомозов О.А., Кузнецов А.Е., Лось В.В., Пресняков О.А. Структурное восстановление изображений, формируемых многоматричными сканирующими датчиками. // *Методы и устройства формирования и обработки сигналов в информационных системах. Межвуз. сб. науч. тр.* Рязань: РГРТА, 2004. С. 88–96.
7. Гомозов О.А., Еремеев В.В., Кузнецов А.Е., Лось В.В., Пресняков О.А., Соловьева К.К. Алгоритмы и технологии обработки информации от КА «Ресурс-ДК». Современные проблемы дистанционного зондирования Земли из космоса. 2008. Т. 5. № 1. С. 69–76.
8. Воронин Е. Г. Метод и результаты геометрической сшивки оптико-электронных космических снимков // Системы наблюдения, мониторинга и дистанционного зондирования Земли: Материалы научн.-техн. конф. - М.: МНТОРЭС им. А. С. Попова, Филиал ФГУП «ГНПРКЦ ЦСКБ-Прогресс» – НПП «ОПТЭК», 2013. С. 256–266.
9. Eremeev V., Kuznetsov A., Myatov G., Presnyakov O., Poshekhonov V., Svetelkin P. Image structure restoration from sputnik with multi-matrix scanners. *Proc. SPIE 9244, Image and Signal Processing for Remote Sensing XX, 92440F* (October 15, 2014); doi:10.1117/12.2066631
10. Еремеев В.В. Методы и информационные технологии межотраслевой обработки многозональных космических изображений: дис. док. техн. наук. – Рязань: РГРТА, 1997. – 312 с.
11. Shewchuk J. What is a good linear finite element? *Interpolation, conditioning, anisotropy and quality measures*, 2003, Technical report, CS, UC Berkeley.

Уважаемые коллеги!

Для тех, кто не успел оформить подписку на второе полугодие 2015 года через ОАО «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.dsra.ru.

Справки по телефонам: (8-495) 362-42-75; (+7 903) 221-79-79 (Алексеева Любовь Ильинична).

E-mail: ntores@mail.ru

УДК 004.932.2

ТЕХНОЛОГИЯ АВТОМАТИЧЕСКОГО КОНТРОЛЯ ТОЧНОСТИ ГЕОПРИВЯЗКИ СПУТНИКОВЫХ ИЗОБРАЖЕНИЙ ПО ОПОРНЫМ СНИМКАМ ОТ КА «LANDSAT-8»

*Кузнецов А.Е., д.т.н., зам. директора НИИ «Фотон» Рязанского государственного радиотехнического университета (РГРТУ), e-mail: foton@rsreu.ru;
Пошехонов В.И., с.н.с. НИИ «Фотон» РГРТУ;
Рыжиков А.С., техник НИИ «Фотон» РГРТУ.*

AUTOMATIC GEOREFERENCE ACCURACY CONTROL TECHNOLOGBASED ON REFERENCE PICTURES FROM THE OBSERVATION SATELLITE «LANDSAT-8»

Kuznetsov A.E., Poshehonov V.I., Ryzgikov A.S.

The technology of automatic control of precision geodetic reference satellite images in the article is described. The technology consists of two parts. First, the creation of a reference data bank. Second, the mechanism of identification of similar objects on the analyzed and the reference image. In the article described the algorithm of a creation of the global bank of reference raster data, based on images of observation satellite «Landsat-8». Also, proposed a reliable and high-performant mechanism of identification of equivalent points on multi-temporal satellite images. This mechanism is based on the comparison of the descriptors of specific fragments.

Keywords: control of precision geodetic reference, raster data bank, points coregistration.

Ключевые слова: контроль точности геопривязки, банк растровых данных, идентификация одноименных точек.

Введение

В настоящее время на центрах приема и обработки спутниковой информации операция контроля точности геопривязки материалов съемки от космических аппаратов высокого пространственного разрешения является обязательной. Это объясняется тем, что по навигационным измерениям положения и ориентации КА не всегда обеспечивается требуемая точность геодезической привязки изображений земной поверхности. Например, при штатной работе звездных датчиков КА «Канопус-В» ошибки координатной привязки маршрутов съемки не превосходят 60 м. Однако встречаются ситуации, когда из-за сильного углового движения или «засветки» астродатчика резко снижается точность определения углового положения спутника, что приводит к ошибкам геопривязки видеоданных от 1.5 – 2 до 4 км. Поэтому для выявления подобных маршрутов необходимо сравнивать геодезические координаты представленных на снимке характерных объектов с опорными значениями.

Операция оценки координатного рассогласования выполняется в ручном режиме с использованием картографических сервисов Google Maps, Яндекс Карты и др., предоставляющих возможность измерения геодезических координат и высоты указываемых оператором на экране объектов. Очевидно, что при ручном контроле каждого маршрута съемки и зависимости этого процесса от Internet-ресурсов нельзя организовать автомати-

Рассмотрена технология автоматического контроля точности геодезической привязки спутниковых изображений, включающая создание опорного банка данных и механизм идентификации одноименных объектов на анализируемом и опорном изображениях. Описывается алгоритм создания глобального банка опорных растровых данных на основе снимков КА «Landsat-8». Предложен надежный и высокопроизводительный механизм идентификации одноименных точек на разновременных спутниковых изображениях, основанный на сопоставлении дескрипторов характерных фрагментов.

ческую обработку и каталогизацию видеоданных, что очень важно для достижения высокой производительности наземного комплекса обработки и оперативного обслуживания потребителей. Поэтому целью работы является разработка технологии автоматического контроля точности геопривязки спутниковых изображений, не зависящей от внешних ресурсов. Поставленная цель достигается решением двух задач: созданием опорного растрового банка данных на заданную территорию и разработки высоконадёжного и высокопроизводительного механизма поиска одноименных точек на анализируемом и опорном изображениях.

Проектирование опорного банка данных

В работе [1] рассматриваются вопросы уточнения геопривязки изображений от китайского спутника по данным от КА «Landsat-8». Геокодированные снимки от этого КА находятся в открытом доступе по адресу <http://earthexplorer.usgs.gov/>. Точность геокодирования панхроматических изображений с пространственным разрешением 15 м по критерию CE 90 составляет не менее 12 м. Изображения представлены отдельными

кадрами размером 190x190 – 195x195 км. Поэтому информацию от этого спутника будем использовать для создания опорного банка данных.

К сожалению, на базе отдельных кадров крайне затруднительно организовать высокоскоростной поиск одноименных сюжетов на анализируемом и опорном изображениях. Поэтому, по аналогии с картографическими сервисами, создадим многомасштабное непрерывное растровое покрытие опорной территории земной поверхности. Для этого трансформируем каждый кадр от КА «Landsat-8» из проекции UTM в непрерывную проекцию Меркатора, используя уравнения обратного геометрического преобразования,

$$x' = F(x, y), y' = \Phi(x, y), \quad (1)$$

где (x', y') и (x, y) – планарные координаты пикселей в проекции UTM и Меркатора соответственно; (F, Φ) – уравнения, описывающие пересчет координат пикселей из проекции Меркатора в проекцию UTM. Для проекции Меркатора, в отличие от большинства картографических Web-сервисов, используем точные уравнения картографического проектирования,

$$x = a \cdot \lambda,$$

$$y = a \cdot \ln \left[\tan \left(\frac{\pi}{4} + \frac{\phi}{2} \right) \cdot \left(\frac{1 - e \cdot \sin(\phi)}{1 + e \cdot \sin(\phi)} \right)^{e/2} \right], \quad (2)$$

где λ, ϕ – долгота и широта пикселя с координатами (x, y) ; a, e – значения большой полуоси и эксцентриситета эллипсоида.

Для организации многомасштабного пирамидального представления опорного изображения $B_i, i = \overline{0, I}$, где i – номер слоя в пирамиде, трансформирование будем выполнять с разными значениями параметра a ,

$$a = a_i = \frac{2\pi}{L_i}, \quad (3)$$

где величина L_i соответствует размеру изображения B_i и определяется соотношением

$$L_i = 2^{I-i} \cdot L_{\min}, \quad (4)$$

где $L_{\min} = 256$.

Пространственное разрешение слоя i на широте ϕ

$$r_{i\phi} = \frac{\cos(\phi)L_e}{L_i}, \quad (5)$$

где L_e – длина экватора в метрах. При добавлении в банк данных снимка осуществляется поиск слоя i , для которого выполняется соотношение a , где ϕ' – широта центра снимка, R – его разрешение. Далее с использованием уравнений (2) растровые кадры трансформируются в проекцию Меркатора и формируют изображение B_i .

Слои пирамиды представляют собой изображения значительного размера. Поэтому для работы с такими растрами используем механизм тайловой организации данных. Непрерывные изображения хранятся в виде тайлов – квадратных растров размером $L_{\min} \times L_{\min}$. Количество тайлов, из которых состоит слой i , определяется значением 4^{I-i} , а сторона слоя состоит из $T_i = 2^{I-i}$ тайлов. По заданным индексам тайла достаточно просто найти соответствующий фрагмент изображения B_i .

Обозначим $g = \overline{0, T_i}, k = \overline{0, T_i}$ – индексы тайла по осям x, y соответственно, (m, n) – координаты точки внутри тайла. Тогда точка $O'(m, n)$ тайла $D_{k,g}$ эквивалентна точке $O'(x_0, y_0)$ виртуального изображения B_i , где

$$\begin{aligned} x_0 &= g \cdot L_{\min} + n, \\ y_0 &= k \cdot L_{\min} + m. \end{aligned} \quad (6)$$

Обратное отображение так же элементарно:

$$k = \text{ent} \left[\frac{y_0}{L_{\min}} \right], k = \text{ent} \left[\frac{x_0}{L_{\min}} \right], \quad (7)$$

$$n = x_0 \bmod (L_{\min}), m = y_0 \bmod (L_{\min}),$$

где ent – оператор взятия целой части.

Таким образом, при извлечении фрагментов непрерывного изображения используются только те тайлы, которые необходимы в данный момент, что позволяет избежать «холостых» обращений к дисковой памяти, уменьшить число кэш-промахов и увеличить скорость выдачи опорного растра.

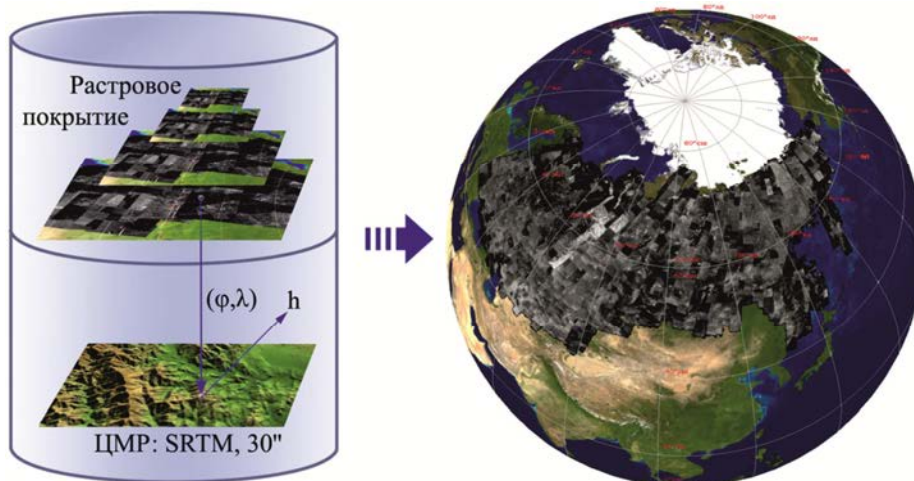


Рис.1. Структура банка опорной информации и пример растрового покрытия территории РФ данными от КА «Landsat-8»

На заключительном этапе формирования опорного банка данных подключим к растровому покрытию B_i цифровую модель рельефа SRTM (данные находятся в открытом доступе по адресу <http://srtm.csi.cgiar.org>) с тем, чтобы обеспечить возможность измерения в опорной точке не только геодезических координат, но и высот объектов. На рис. 1 показана структура банка опорной информации и растровое покрытие, сформированное на территории РФ и стран СНГ.

Алгоритм автоматического обнаружения одноименных точек на опорном и анализируемом изображениях

Задача обнаружения одноименных точек и вычисления их координатных рассогласований имеет ряд особенностей.

Во-первых, непрерывное покрытие B_i синтезировано из одновременных снимков, которые могут существенно отличаться по текстуре от анализируемого изображения B' , что делает крайне ненадежной работу процедуры площадной корреляции. Пример таких снимков изображен на рис. 2 (пунктиром выделены одноименные области с различающейся текстурой).

Во-вторых, анализируемые изображения B' представляют собой массивы сверхбольшого объема (от нескольких десятков до сотен гигабайт), которые могут иметь координатные рассогласования с опорным сним-

ком до нескольких сотен пикселей. При этом оценка этого рассогласования должна выполняться быстро и практически со 100 % надежностью.

Исходя из перечисленных особенностей, проанализируем возможные решения по поиску на снимках одноименных точек. Так, в работе [2] для учета значительных взаимных координатных рассогласований поиск одноименных объектов выполняется по пирамиде изображений B_i' , B_i . Для исключения влияния текстуры сюжетов изображения предварительно приводят к контурному виду, а результаты корреляционного сопоставления анализируются с помощью целой группы статистических проверок вида корреляционной функции. К сожалению, этот подход для случаев, показанных на рис. 2, из-за низкого коэффициента корреляции приводит к большой доле ложных сопоставлений.

Для повышения надежности площадной корреляции, во-первых, будем анализировать фрагменты изображений B_i' , содержащие характерные объекты. В качестве детектора характерных объектов будем использовать детектор Харриса [3]. Этот оператор позволяет находить на изображениях характерные точки, в окрестности которых у градиента изображения имеются два ярко выраженных доминирующих направления. Для их обнаружения выполняется последовательность шагов.

Сгладим изображение B' с использованием фильтра Гаусса с окном 3×3 .

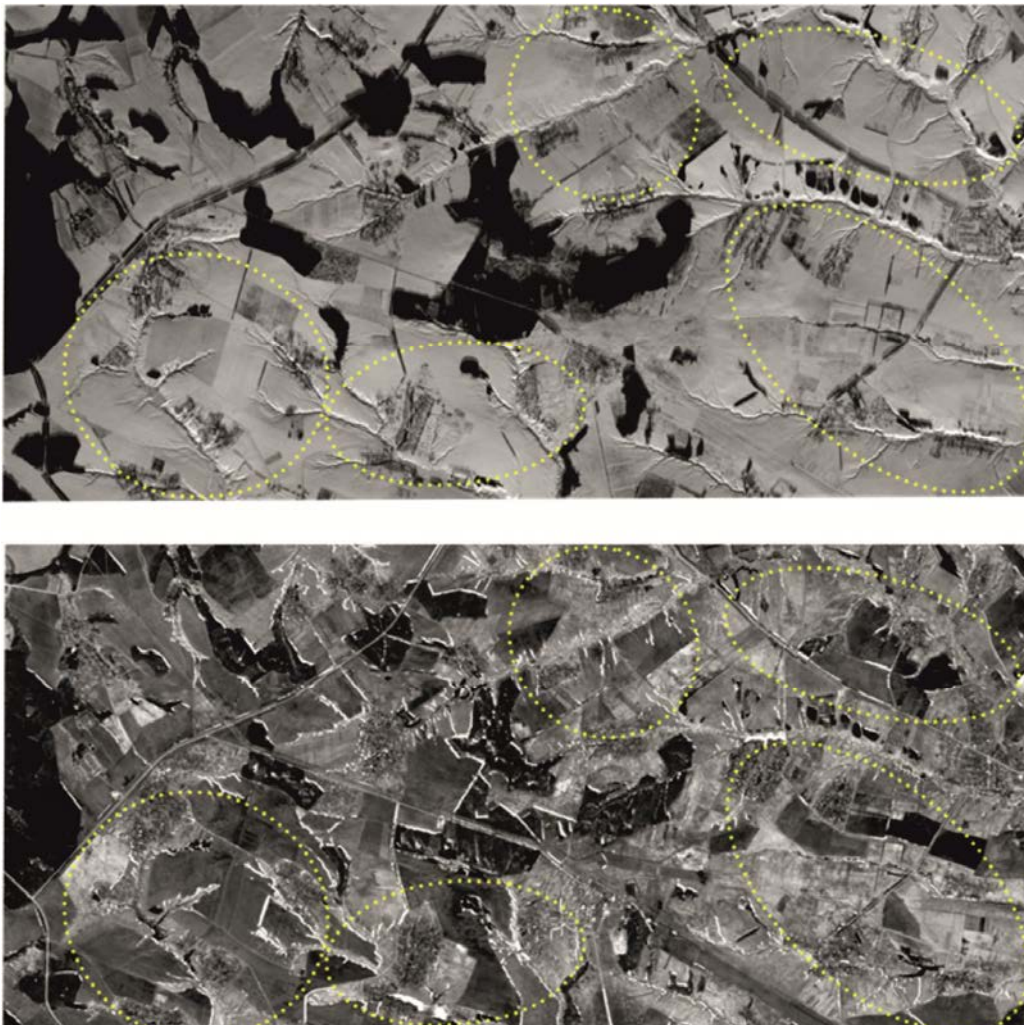


Рис. 2. Изображения одновременных снимков от КА «Канопус-В» и «Landsat-8»

Продифференцируем полученное изображение путем свертки с окнами $(-1 \ 0 \ 1)$, $(-1 \ 0 \ 1)^T$. На этом шаге формируются массивы производных B' , $\partial B'_x = \{\partial b'_x(x, y)\}$, $\partial B'_y = \{\partial b'_y(x, y)\}$ по осям x и y соответственно.

Сформируем массивы V_x , V_y , V_{xy} , характеризующие скорость изменения градиента по горизонтали, вертикали и диагонали соответственно:

$$\begin{aligned} V_x &= \{v_x(x, y)\} = \partial b'_x(x, y) \cdot \partial b'_x(x, y), \\ V_y &= \{v_y(x, y)\} = \partial b'_y(x, y) \cdot \partial b'_y(x, y), \\ V_{xy} &= \{v_{xy}(x, y)\} = \partial b'_x(x, y) \cdot \partial b'_y(x, y). \end{aligned} \quad (8)$$

Сгладим массивы V_x , V_y , V_{xy} фильтром Гаусса с окном $S \times S$, причем размер S задает масштаб, в котором будут определены характерные точки.

Сформируем симметричные матрицы производных для каждой точки (x, y) анализируемого изображения,

$$\mathbf{I} = \begin{bmatrix} v_x(x, y) & v_{xy}(x, y) \\ v_{xy}(x, y) & v_y(x, y) \end{bmatrix}. \quad (9)$$

Характерные точки имеют значительную величину собственных значений матрицы \mathbf{I} . После выполнения последнего шага строится множество координат характерных точек, для которых значение

$$R(x, y) = \det \mathbf{I} - \mathcal{G} \cdot \text{trace } \mathbf{I} \quad (10)$$

превышает заданный порог, где $\mathcal{G} = 0.04$ – эмпирический коэффициент.

Во-вторых, для отбраковки ложных сопоставлений будем использовать метод обратного поиска. Его суть состоит в следующем: после нахождения образа фрагмента T' на изображении B_i , осуществляется поиск образа фрагмента T на изображении B_i' . Если смещения, найденные при прямом и обратном поиске, отличаются не более чем на 1 пиксель, то отождествление фрагментов T и T' считаем достоверным.

Исследования показали, что с использованием предложенных решений доля ложных сопоставлений уменьшается и они носят случайный характер, а истинные координатные рассогласования обладают систематичностью. Поэтому для определения достоверных координатных рассогласований изображений B_i' и B_i , характеризуемых вектором $\mathbf{r}_i = (\Delta x_i, \Delta y_i)$, сформируем гистограмму распределения этого параметра $G(\mathbf{r})$. Откуда искомое координатное рассогласование \mathbf{r}_0 определяется как $\mathbf{r}_0 = \arg \text{med} G(\mathbf{r})$.

К сожалению, временные затраты рассмотренного подхода достаточно велики и в отдельных случаях оператор быстрее находит одноименные точки, чем процедура автоматического поиска. Поэтому была исследована возможность поиска положения различных фрагментов в соответствующих им зонах поиска на отдельных ядрах CPU. Так как распределение характерных фрагментов неравномерно и зависит от сюжета изображения, соответствующие им зоны поиска могут нахо-

диться в различных частях опорного изображения. На практике это означает необходимость работы со значительно большим, по сравнению с обработкой одного фрагмента, объемом данных. При превышении объема кэш-памяти верхнего уровня возникают систематические кэш-про-махи, связанные с одновременными операциями доступа к различным участкам ОП. Вследствие конкурентного доступа к памяти ускорение за счет использования нескольких ядер системы несущественно, а значительное превышение объема рабочего набора данных над объемом кэша приводит к замедлению работы по сравнению с однопоточной реализацией.

Значительно более эффективным с точки зрения распараллеливания представляется метод поиска одноименных точек SURF, основанный на извлечении из сопоставляемых изображений B' и B_0 компактных описаний характерных фрагментов – дескрипторов и дальнейшим их сопоставлении между собой.

В алгоритме SURF для обнаружения характерных точек осуществляется поиск не углов, а пятен (blobs), т.к. дескрипторы такого рода фрагментов удается сопоставить с большей степенью достоверности. Для нахождения характерных точек изображения B выполняется последовательность шагов:

Формирование матриц $D_\tau = [d_\tau(x, y)]$, где $d_\tau(x, y)$ – определитель матрицы Гессе. Для вычисления вторых производных изображения B в точке (x, y) применяется бинаризованная аппроксимация фильтра Лапласиана гауссиана с размером окна фильтра τ , $\tau = \{9, 15, 21, 27\}$ [4]. Локальные экстремумы функции $d_\tau(x, y)$ соответствуют точкам максимального изменения яркости – пятнам и углам с размером, определяемым параметром τ .

Определение координат характерных точек. Точка с координатами (x, y) является характерной точкой изображения B , если $d_\tau(x, y)$ является локальным максимумом в окне $3 \times 3 \times 3$, т.е. $d_\tau(x, y)$ больше гессиан соседних пикселей, построенных с размером окна фильтра $\{\tau - 6, \tau, \tau + 6\}$.

Субпиксельное уточнение положений характерных точек путем интерполяции найденных гессиан.

Для построения дескрипторов характерных точек для каждой точки выполняются следующие действия:

Определение ориентации дескриптора. Для этого с использованием фильтра Хаара вычисляются точечные градиенты в окрестности радиусом 6 отсчетов. Направление дескриптора формируется путем усреднения взвешенных значений вейвлета Хаара.

Вычисление дескриптора особой точки. Для вычисления дескриптора формируется прямоугольная окрестность характерной точки со стороны прямоугольника 40 . Эта окрестность разбивается на 16 квадратов, в каждом из которых для всех точек в регулярной сетке 5×5 с помощью фильтра Хаара вычисляется градиент. Дескриптор формируется из угла ориентации градиента, знака следа матрицы Гессе и четырех интегральных характеристик каждого квадрата – взвешенных сумм значений вейвлета Хаара по осям x и y , взвешенных сумм модулей этих значений. Вес определяется как значение гаус-

сиана с центром в особой точке. В соответствие с методикой авторов SURF, сформированные направления градиентов «поворачиваются» в соответствии с доминирующим направлением, что гарантирует инвариантность дескрипторов к повороту снимка.

Для определения одноименных точек согласно оригинальной методике выполняется сопоставление всех дескрипторов одного снимка с дескрипторами другого, причем мерой близости является Евклидово расстояние, рассчитываемое по всем компонентам дескриптора. Для ускорения операции сопоставления предлагается использовать знак следа матрицы Гессе (очевидно, что для светлых фрагментов на темном фоне он отрицателен и наоборот).

В ходе экспериментальных исследований установлено, что около 90 % времени на отождествление занимает сопоставление дескрипторов. Поэтому для сокращения временных затрат модифицируем алгоритм сопоставления дескрипторов.

Заметим, что в рамках решаемой задачи после трансформации в единую картографическую проекцию сопоставляемые снимки ориентированы в одном направлении и характеризуются схожим масштабом. Дескрипторы SURF, не нормированные в соответствии с доминирующим направлением, устойчивы к повороту до 15 градусов. В соответствии с этим предлагается для сопоставления множеств дескрипторов снимка Q' и опорного изображения Q следующий алгоритм.

Выделить непересекающиеся подмножества Q_τ , Q'_τ , $\tau = \{9, 15, 21, 27\}$ так, чтобы дескрипторы каждого из подмножеств описывали фрагмент, являющийся характерным в масштабе, определяемом параметром τ .

Выделить из каждого Q_τ , Q'_τ пересекающиеся подмножества $Q_{\tau\alpha_j}$, $Q'_{\tau\alpha_j}$ такие, что все дескрипторы множеств $Q_{\tau\alpha_j}$, $Q'_{\tau\alpha_j}$ характеризуются углом доминирующего направления $\alpha \in [a_j, a_{j+1} - \beta]$, где параметр a_j определен рекурсивно,

$$a_j = 0, a_{j+1} = a_j + \delta; \tag{11}$$

величина $\delta - \beta$ определяет размер диапазона пересечения.

Осуществить сопоставление дескрипторов всех соответствующих множеств $Q_{\tau\alpha_j}$, $Q'_{\tau\alpha_j}$.

Однако наибольший прирост производительности дает параллельное вычисление компонент дескриптора. Для определения направления дескриптора и формирования его компонент необходимо вычислить 25 фильтров Хаара в 16 квадратах, что составляет 400 фильтров, в то время как фрагмент изображения размером 40x40 пикселей, необходимый для расчета дескриптора, занимает 1600 байт. С учетом малого размера блока данных целесообразно параллельно рассчитывать фильтры Хаара для n пикселей дескриптора, где число n определяется исходя из количества ядер системы и наличия технологий типа hyper-threading. При реализации на современном GPU все фильтры Хаара рассчитываются параллельно и с использованием быстродействующей разделяемой между потоками shared memory. Кроме того, в случае наличия большого числа ядер общего назначения, возможно параллельное вычисление нескольких дескрипторов без потери производительности (схема доступа к фрагментам снимка в памяти изображена на рис. 3).

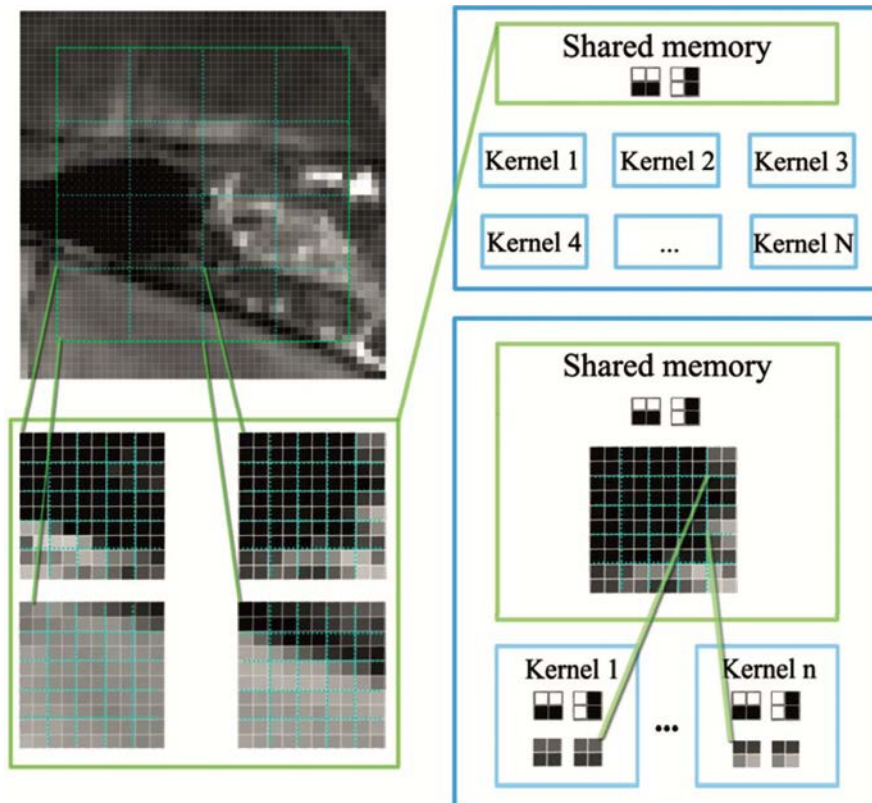


Рис. 3. Схема формирования дескриптора на GPU

Для сравнения, вычисление корреляционной функции двух фрагментов размером 128x128 пикселей без учета буфера на запись результатов корреляции требует по меньшей мере 32768 байт и не может быть эффективно реализовано с использованием shared-memory (стандартный размер которой составляет 16384 байт). Помимо этого, реализовано параллельное формирование матриц Гессе, собственно обнаружение характерных точек путем сопоставления с соседями. Так как в ходе этих операций обрабатывается значительный объем данных, используется скользящее окно небольшого размера, фрагменты внутри которого параллельно используются в нескольких потоках. В результате в зависимости от быстродействия оперативной памяти достигается 2-4 кратный прирост производительности относительно однопоточного варианта.

По итогам работы выполнена многопоточная реализация алгоритма поиска одноименных точек на базе SURF-дескрипторов на CPU (Intel Xeon 2.40GHz x 36) и GPU (NVIDIA Tesla K20c) с использованием примитивов библиотеки OpenCV (<http://opencv.org/>). Поиск одноименных точек на снимке размером 1920x985 пикселей и изображении размером 3840x3840 пикселей занимает 50 – 110 секунд для CPU реализации и до 4 секунд для GPU.

Заключение

Таким образом, по результатам проведенных исследований получены следующие результаты.

Создан банк опорной растровой информации на территорию РФ и страны СНГ. Растровое покрытие этой

территории организовано в виде пирамиды разномасштабных слоев с разбивкой на тайлы и занимает объем 300 Гб. Такое представление информации позволяет реализовать доступ к любому участку покрытия в режиме реального времени.

На базе алгоритма SURF разработан высокопроизводительный и надежный механизм идентификации координат одноименных объектов на опорном и анализируемом разновременных изображениях, имеющих текстурные различия и содержащих облачные объекты. Найденные координаты опорных точек и их высоты передаются в процедуру оценки точности геопривязки маршрутов съемки.

Созданная технология автоматического контроля геопривязки адаптирована к информации от КА «Канопус-В» и проходит в настоящее время отработку в Научном центре оперативного мониторинга Земли.

Литература

1. Chabita Devaraj, Chintan A. Automated geometric correction of multispectral images from High Resolution CCD Camera (HRCC) on-board CBERS-2 and CBERS-2B// ISPRS Journal of Photogrammetry and Remote Sensing, 2014, 89.
2. Кузнецов А.Е., Светелкин П.Н. Формирование цветных снимков по спутниковым изображениям среднего и высокого пространственного разрешения // Цифровая обработка сигналов, 2009, №3, С. 36-40.
3. Гонсалес Р., Вудс Р., Цифровая обработка изображений. – Техносфера, 2000, 1072с.
4. Herbert Bay, Tinne Tuytelaars, and Luc Van Gool, «Speeded Up Robust Features», ETH Zurich, Katholieke Universiteit Leuven.

ПРИГЛАШЕНИЕ

на XI Международную научную конференцию

Перспективные технологии в средствах передачи информации – ПТСПИ-2015

12 – 14 ноября 2015 г. Владимир-Суздаль, Россия

ОРГАНИЗАТОРЫ:

Министерство образования и науки РФ, Российское НТО РЭС им. А.С. Попова, Международная академия связи, Институт радиотехники и электроники им. В.А. Котельникова РАН, Владимирский государственный университет имени А.Г. и Н.Г. Столетовых, Университет Фридриха Александра (Германия); Фраунгоферовский институт интегральных схем (Германия); Фирма National Instruments (США).

ТЕМАТИЧЕСКИЕ РАЗДЕЛЫ КОНФЕРЕНЦИИ:

1. Технологии и сети инфокоммуникаций;
2. Методы обработки информации;
3. Помехозащищенность инфокоммуникационных систем;
4. Моделирование и автоматизированное проектирование инфокоммуникационных систем и устройств.
5. Системы и комплексы инфокоммуникаций;
6. Молодежная научная школа «Интеллектуальные системы и технологии передачи информации»;
7. Круглый стол: «Перспективы развития инфокоммуникаций».

СПРАВКИ:

Сайт: <http://ptmm.vlsu.ru>

Телефоны: (4922) 479-960; 479-739; Факс: (4922) 479-960

E-mail: ptmm@vlsu.ru

Адрес: ул. Горького, д. 87, ВлГУ, ФРЭМТ, г. Владимир, 600000.

УДК 004.932.4

ОБНАРУЖЕНИЕ ИГРОВЫХ ЭПИЗОДОВ В ВИДЕОПОТОКЕ

*Петрова К.Ю., к.т.н., руководитель проектов, Исследовательский Центр Самсунг, Москва,
e-mail: xenyu.petrova@samsung.com;*

*Рычагов М.Н., д.ф.-м.н., проф., начальник управления, Исследовательский Центр Самсунг, Москва,
e-mail: michael.rychagov@samsung.com;*

*Седунов С.М., старший инженер, Научно-Исследовательский Центр Хуавей, Москва,
e-mail: Sedunov.Sergey@huawei.com.*

FIELD GAME EPISODE DETECTION IN VIDEO SEQUENCES

Petrova X.Y., Rychagov M.N., Sedunov S.M.

Description of classifier implementing field game episode detection in video sequences in real time is provided. Method is based on visual cues, described by one and multi-dimensional distributions of color and texture features. Classifier has form of directed acyclic graph with threshold transforms, linear sub-classifiers and Boolean logical functions in the nodes. Solution can be used in TV receivers for automatic selection of settings, providing comfortable viewing of sport programs and as a pre-processing classification cascade in off-line solutions for high precision video classification.

Keywords: video classification, real time, sport, soccer, visual cue, acyclic directed graph, color, texture.

Ключевые слова: классификация видео, реальное время, спорт, футбол, визуальный ключ, ациклический направленный граф, цвет, текстура.

Введение

В современных телевизионных приёмниках принято использовать различные режимы обработки видеосигнала, соответствующего производственным жанрам видео, таким как спорт, кино, музыка и т.п. Например, для обеспечения лучшего качества изображения при приеме спортивных передач обычно рекомендуется выбирать средний контраст, небольшой коэффициент улучшения деталей и низкий коэффициент шумоподавления, в то время как при просмотре фильмов и музыкальных передач эти коэффициенты могут принимать повышенные значения. Таким образом, детектирование игровых эпизодов в реальном времени может быть полезно для автоматической подстройки коэффициентов, обеспечивающей наиболее комфортный просмотр спортивных телепередач.

Основной целью исследований, результаты которых представлены в статье, является разработка алгоритма обнаружения эпизодов, содержащих игровые виды спорта в видеопоследовательности в реальном времени с

Классификация как статических, так и видеоизображений является бурно развивающейся областью и обладает большим потенциалом коммерциализации. В настоящей работе решается задача определения игровых эпизодов в видеопотоке на основе применения т.н. визуальных ключей, характеризующихся одномерными и двумерными распределениями цветовых и текстурных признаков. Реализована технология обработки видео, обеспечивающая требуемую классификацию с точностью до одного кадра в реальном времени и позволяющая простую аппаратную реализацию.

целью последующей автоматической подстройки настроек изображения в ТВ-приемнике (рис. 1).

Обзор работ по классификации видеопоследовательностей

Известные работы по классификации видеопоследовательностей можно разделить по следующим признакам: цель классификации (например, определение жанра или детектирование определенных объектов), используемые модальности (такие как видео, аудио, субтитры), выбор признаков и тип классификатора. Наиболее близкими по целевому признаку являются широко представленные работы по автоматическому определению жанра [1-15].

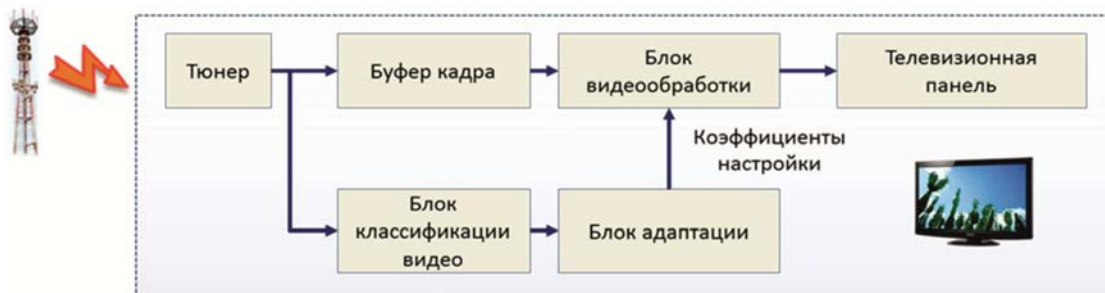



Рис. 1. Распознавание жанра видео в конвейере видеобработки телевизионного приемника



Определение жанра может производиться с использованием различных модальностей, например, субтитров [1], звука [2-4], или видео (как во всех прочих источниках), или одновременно нескольких модальностей [5].

В [1] в качестве признаков использованы субтитры и коэффициенты ДКТ разложения видеок кадров. Авторы получили высокую точность детектирования, однако отметили, что субтитры отсутствуют во многих телевизионных трансляциях. Хотя в некоторых странах, как, например, в США законодатель обязывает вещателей сопровождать теле-трансляции субтитрами. При этом, субтитры не являются описанием того, что демонстрируется на экране, и не генерируются для сцен, в которых отсутствуют диалоги. Наконец, обучение и классификация по такому признаку могут сопровождаться большой вычислительной сложностью, поскольку вектор признаков может состоять из многих тысяч элементов.

В [2] видеопоследовательности, содержащие запись спортивных соревнований, проиндексированы с помощью SVM-классификатора на основании признаков аудиопотока.

Подходы, основанные на анализе аудио потока или субтитров, неприменимы в телевизионном приемнике для автоматической подстройки коэффициентов видеопереработки, поскольку эти признаки не всегда совпадают по времени с соответствующим видеорядом.

Далее будут рассматриваться методы, основанные исключительно на анализе видео-потока.

В [6] производится определение следующих жанров: мультфильмы, реклама, новости и спорт. В качестве базового алгоритма выбран метод опорных векторов, использующий ориентированный ациклический граф (DAGSVM). При этом выделялись 15 визуальных признаков четырех типов (монтаж, цвет, текстура и движение). С точки зрения монтажа оценивалась частота смены эпизодов и соотношение между количеством резких и плавных переходов между сценами. В качестве цветовых признаков выступали гистограмма, средняя яркость и насыщенность, а также процент пикселей с яркостью и насыщенностью выше предопределенного порога. Текстурные признаки соответствовали статистическим свойствам полутонной матрицы смежности: контрасту, однородности, энергии, энтропии и корреляции. Признаки движения определялись как среднее изменение яркости и средняя разность в цветовом пространстве RGB между соседними кадрами, а также доля кадров, которые отличаются незначительно от предыдущего (доля медленных и статических сцен).

В работе [7] проводится классификация кинофильмов по жанрам: боевики, драмы и триллеры. Использовались шесть глобальных признаков: средняя длина эпизода, вариация цвета, движение, освещение (наличие высвечивающего фонаря), статистика типов переходов между сценами. Классификатор реализован в виде дерева принятия решений.

В работе [8] описано применение свойств границ и классификатора K средних для того, чтобы отличить бадминтон, баскетбол, футбол, коньки и теннис. В более ранней работе тех же авторов свойства границ использовались для того, чтобы выделить кадры, содержащие общий план зрительских трибун и рекламу вне игрового поля.

В [9] анализировались траектории лиц и блоков текста для того, чтобы отличить рекламу, новости, комедийные сериалы и мыльные оперы. Классификация была произведена путем нахождения максимальной проекции распределения траекторий классифицируемого фрагмента на множество траекторий обучающего множества.

В [10] кадры видеозаписи лекций разделены на 4 класса. Для этого предлагается квази-оптимальная процедура выбора признаков (из 300 исходных признаков).

В [11-12] посредством анализа движения камеры производилась классификация пяти видов спорта (сумо, теннис, бейсбол, футбол и американский футбол). Для каждого из этих видов спорта характерен особый способ монтажа. Был проведен анализ использования съемки с проводкой, наездов камеры, эпизодов со статичной и трясущейся камерой, а также типов переходов от одного вида движения камеры к другому. Поскольку метод не опирается на цветовую информацию, он может быть очень эффективен при классификации трансляций игр разных спортивных лиг (например, на открытом чемпионате Великобритании по теннису корты – зеленые, а на открытом чемпионате Франции – красные).

В [13] предложен новый низкоуровневый признак – поток энергии активности. Он подавался на вход сети радиальных базисных функций для определения спорта, кино, ландшафтных съемок и новостей. Поток энергии активности измерялся в пределах одного эпизода и был основан на статистике макро-блоков сжатого видео, включая количество I -блоков, блоков с надежно предсказанным движением и блоков с ненадежно предсказанным движением (например, в случае однородного фона при движущейся камере).

В [14] для определения видов спорта (теннис, трековые виды спорта, плавание, яхтинг и велогонки) использовалась концепция «визуальных ключей». Эти ключи придавали семантическое значение низкоуровневым признакам кадров видеопоследовательности. В общей сложности выделялось 17 типов ключей: трек для занятий легкой атлетикой, боксерский ринг, крытый велотрек, океан, лыжный трамплин, бассейн, теннисная сетка, трава, синее небо, открытый велотрек, океан (средний план), беговая дорожка (дальний план), беговая дорожка (средний план), беговая дорожка (крупный план), бассейн (крупный план), теннисный корт. Первые 9 детекторов были обученными нейронными сетями, а последние 8 использовали т.н. текстурные коды. Результаты детектирования визуальных ключей подавались на вход классификатора K средних, который определял, к какому виду спорта относится эпизод. В [15] набор признаков был расширен посредством добавления к нему мультимодальных элементов для детектирования визуальных ключей. При этом применялись скрытые марковские модели для анализа временных зависимостей.

В [16] рассматривалась задача выделения 7 видов спорта (альпинизм, баскетбол, автогонки, гольф, прыжки с трамплина, футбол и мотогонки) и 5 типов сюжетов, имеющих отношение к спорту: финальные титры, комментаторы в студии, начальные титры, трейлер и турнирная таблица. Решение принималось на основе вычисления векторов цветовой когерентности [17].

В [18] произведена классификация 7 популярных видео жанров: мультфильмы, реклама, документальные фильмы, кино, музыкальные клипы, новости и спорт. Было использовано три категории дескрипторов – на уровне цвета, динамики и структуры. Цветовые свойства описывались глобально посредством статистики распределения различных цветов, элементарных цветов, цветовых свойств (яркость, насыщенность) и отношений между ними. С точки зрения динамики, оценивался ритм видео, статистика движений и процентная доля плавных переходов. Структурная информация извлекалась на уровне кадра путем построения гистограмм контуров и выявления отношений между ними.

Инструментарий, используемый для классификации изображений и видео весьма разнообразен. Ранее были продемонстрированы эффективность метода решающие деревья [7], анализа основных компонентов [19], метода опорных векторов [2, 3], нейронных сетей [5, 14], сетей Кохонена [20], скрытых марковских моделей [21]. В [22] использован классификатор по принципу ближайшего соседа. Некоторые исследователи используют случайный лес или байесовский подход [23].

В [20] сделана интересная попытка установить соответствие между семантической концепцией и списком синонимов этой концепции, найденных в WordNet. Эта идея может послужить основой для быстрого и эффективного синтеза визуального классификатора на основе вербального описания, сделанного экспертом. Похожая идея описана в [24], где банк классификаторов визуальных объектов использует набор примитивных классификаторов, которые можно использовать в качестве входов более сложных классификаторов.

С точки зрения решаемой задачи интерес представляют также методы определения смены сцены, однако мы ограничились простейшим случаем – прямым переходом, поскольку это преобладающий тип перехода в спортивных трансляциях.

Требования к алгоритму классификации видеок кадров в реальном времени

Будем предъявлять следующие требования к алгоритму классификации:

1. Для управления настройками конвейера видеопработки результат классификации должен быть получен для каждого отдельного кадра и не иметь резких скачков за исключением случая смены сцены. В связи с этим должен использоваться лишь один проход по видеоккадру.

2. Алгоритм должен быть универсальным, т.е. игровые эпизоды должны быть дифференцированы от всех прочих типов сцен, включая кино, новости, мультфильмы, компьютерную графику, концерты и т.д.

3. Алгоритм должен быть нечувствителен к качеству видео потока (поддерживать различное разрешение, как стандартное, так и высокое, быть нечувствительным к различным методам и качеству сжатия).

4. Сложность алгоритма должна быть невысокой, допускающей аппаратную реализацию в виде части микросхемы видеопработки на чипе. Поэтому желательно, чтобы при анализе пикселей использовалась ограниченная локальная окрестность.

5. Алгоритм должен иметь модульную организацию для обеспечения его взаимодействия с другими подсистемами видеопработки.

Описание алгоритма

Синтез алгоритма сопровождался формированием базы данных, состоящей из видеофрагментов различного содержания, разрешения и качества сжатия. Анализ этих данных позволяет сделать следующие выводы:

1. Важность получения правильного результата не одинакова для различных кадров: более важно получить стабильный результат детектирования для кадров, содержащих зеленое поле с движущимися по нему игроками, чем для кадров со зрительскими трибунами или комментатором, или даже для игрока, демонстрируемого крупным планом.

2. Не все ошибки классификации имеют одинаковый вес. Так, вполне допустимо перепутать футбол с бейсболом, но совершенно недопустимо перепутать футбол с музыкальным клипом.

3. Видеоданные (на уровне отдельных кадров) по своей природе обладают очень большим смещением, так как распределение больше всего зависит от длительности эпизодов, содержащих кадры с близкими характеристиками. Т.е. частота довольно слабо связана с важностью.

Чтобы формализовать эти требования, вводится дополнительная под-классификация в виде простой иерархии и весовых коэффициентов для матрицы несоответствий. Матрица несоответствий – это популярный инструмент для оценки качества алгоритмов классификации [25]. Категорию «игровые виды спорта» мы подразделили на три подкатегории: С1 (игровые эпизоды, футбол), С2 (игровые эпизоды, игры на поле кроме футбола), С3 (спорт, неигровые эпизоды). В рамках этой классификации в категорию С4 попали все остальные типы кадров. Весовые коэффициенты для матрицы несоответствий показаны в табл. 1.

Для упрощения синтеза алгоритма, класс С1 был подразделен на три класса: дальний план, общий план и крупный план. Для каждого из классов строился отдельный классификатор. На первом шаге рассматривался наиболее простой и очевидный случай: для того, чтобы отличить кадры дальнего и общего планов демонстрации футбольного матча от кадров прочих жанров, в качестве признака использовалось процентное соотношение зеленых пикселей. Дальнейший синтез проводился с использованием усеченной базы данных, состоящей только из кадров, на которых классификатор, синтезированный на первом шаге, выдает ошибку первого рода.

Таблица 1.

		Спорт			Не-спорт
		C1	C2	C3	
Спорт	C1	0	1/100	1/10	1
	C2	1/100	0	1/10	1
	C3	1/10	1/10	0	1/10
Не спорт	C4	1	1	1/100	0

За основу берется формула $Gr0 = G > RB_{\max}$, где RB_{\max} – максимум значений красного и синего каналов.

Классификация на основе этой формулы обладает недостатками: 1) почти все белые пиксели, 2) пиксели с очень низкой насыщенностью, 3) очень темные (почти черные) пиксели, а также 4) синеватые и 5) желтые пиксели классифицируются как зеленые.

Недостатки с 1-го по 4-й были исправлены путем добавления в формулу члена:

$Gr1 = Gr0 \wedge S_{RGB} > 80 \wedge (R+B < \frac{3}{2}G \vee R+B < 255 < R-B < 35) \wedge Y > 80$, где $S_{RGB} = R+G+B$. Пятый недостаток был исправлен благодаря введению дополнительного классификатора желтых пикселей:

$Y_e = B < G \wedge B < R \wedge 9 \cdot (RG_{MAX} - RG_{MIN}) < RG_{MIN} - B \wedge S > 0.2 \wedge Y > 110$, где RG_{MAX} – максимум зеленого канала G и красного канала R , RG_{MIN} – минимум этих двух значений, а S – насыщенность цвета. Окончательный вид формулы для детектирования зеленых пикселей имеет вид $Gr2 = Gr1 \wedge \neg Y_e$.

Определения пропорции зеленых пикселей далеко не достаточно для решения поставленной задачи. Некоторые игровые эпизоды содержат очень небольшое количество зеленых пикселей, так что даже человек может совершить ошибку при отсутствии сопровождающего текста или аудио-комментария. Поэтому используется условие, что кадр с нулевым количеством зеленых пикселей относится к классу S_4 .

В основу классификации прочих видов сцен были положены следующие наблюдения.

Спортивные сцены, в особенности футбол, характеризуются наличием зеленых пикселей высокой насыщенности. При этом аналогичные зеленые пиксели, воспроизводящие изображение газона, обычно имеют очень небольшие значения в синем цветовом канале, относительно высокую яркость и насыщенность. Диапазон изменения яркости зеленых пикселей при этом не широк. Если значения градиента яркости в пределах зеленых областей велики, то, скорее всего, кадры соответствуют изображениям естественных объектов, а не футбольного поля. В игровых видах спорта яркие и насыщенные пятна обычно соответствуют форме игроков, а белые

области – разметке поля и номерам игроков. При съемке крупным планом игрока в кадре обычно присутствует хотя бы небольшое количество зеленых пикселей. Эти наблюдения выражаются следующим эмпирическим соотношением:

$W = S_{RGB} > 384 \wedge \max(R, G, B) - \min(R, G, B) < 30$, где $S_{RGB} = R+G+B$.

Детектор ярких и насыщенных цветов формализуется в виде

$B_s = \max(R, G, B) > 150 \wedge \max(R, G, B) - \min(R, G, B) \geq \max(R, G, B) / 2$.

Детектор тона кожи заимствован из литературы [26]:

$S_k = (G \neq 0) \wedge B \leq G + \frac{G}{2} \wedge S_{RGB} > \frac{267R}{2^7} \wedge B \leq \frac{83 \cdot S_{RGB}}{2^8} \wedge G \leq \frac{83 \cdot S_{RGB}}{2^8}$.

Правила классификации базируются на следующих кросс-доменных признаках (рис. 2): доля зеленых пикселей F_1 ; доля пикселей тона кожи F_2 ; средняя яркость всех пикселей F_3 ; среднее значение градиента для зеленых пикселей F_4 ; доля ярких и насыщенных пикселей F_5 ; средняя насыщенность зеленых пикселей F_6 ; доля белых пикселей F_7 ; средняя яркость зеленых пикселей F_8 ; среднее значение синего цветового канала зеленых пикселей F_9 ; степень компактности гистограммы яркости зеленых пикселей F_{10} .

Доля зеленых пикселей вычисляется по формуле

$F_1 = \frac{\sum_{i=1..w, j=1..h} \delta(G_r(i, j))}{w \cdot h}$, где w – ширина кадра, h – высота кадра, i, j – координаты пикселя, $G_r(i, j)$ – результат работы детектора зеленых пикселей, а δ – функция, преобразующая логический тип в вещественный $\delta(x) = \begin{cases} 0, & \neg x \\ 1, & x \end{cases}$. Доля пикселей тона кожи вычисляется как $F_2 = \frac{\sum_{i=1..w, j=1..h} \delta(S_k(i, j))}{w \cdot h}$, где $S_k(i, j)$ – резуль

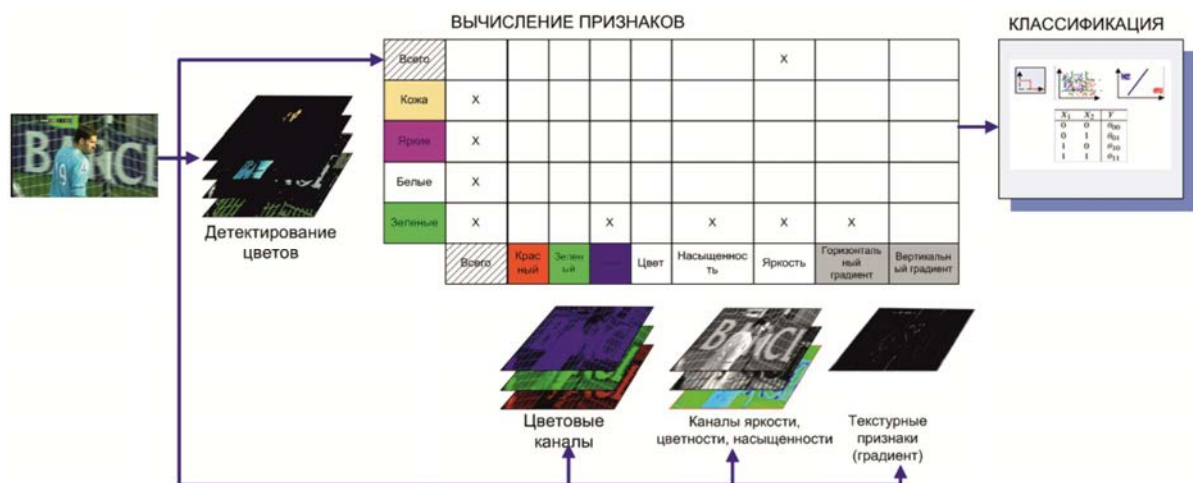


Рис. 2. Вычисление кросс-доменных признаков

тат работы детектора кожи. Средняя яркость вычисляется как $F_3 = \frac{\sum_{i=1..w, j=1..h} Y(i, j)}{w \cdot h}$, где $Y(i, j)$ – яркость пикселя. Среднее значение градиента вычисляется как $F_4 = \frac{\sum_{i=1..w, j=1..h} |D_Y(i, j)| \cdot \delta(G_r(i, j))}{\sum_{i=1..w, j=1..h} \delta(G_r(i, j))}$, где горизонтальная производная яркости D_Y получена путем свертки яркостной компоненты Y с линейным фильтром $K_{grad} = [0 \ 0 \ 0 \ 1 \ 0 \ 0 \ -1]$.

Доля ярких и насыщенных пикселей вычисляется по формуле $F_5 = \frac{\sum_{i=1..w, j=1..h} \delta(B_s(i, j))}{w \cdot h}$, где $B_s(i, j)$ – результат детектирования ярких и насыщенных пикселей. Средняя насыщенность зеленых пикселей определяется в виде $F_6 = \frac{\sum_{i=1..w, j=1..h} S(i, j) \cdot \delta(G_r(i, j))}{\sum_{i=1..w, j=1..h} \delta(G_r(i, j))}$, где $S(i, j)$ – насыщенность. Доля белых пикселей вычисляется как $F_7 = \frac{\sum_{i=1..w, j=1..h} \delta(W(i, j))}{w \cdot h}$, где $W(i, j)$ результат детектирования белых пикселей. Средняя яркость зеленых пикселей вычисляется как

$F_8 = \frac{\sum_{i=1..w, j=1..h} Y(i, j) \cdot \delta(G_r(i, j))}{\sum_{i=1..w, j=1..h} \delta(G_r(i, j))}$. Среднее значение синего цветового канала зеленых пикселей вычисляется по формуле: $F_9 = \frac{\sum_{i=1..w, j=1..h} B(i, j) \cdot \delta(G_r(i, j))}{\sum_{i=1..w, j=1..h} \delta(G_r(i, j))}$. Степень компактности гистограммы яркости зеленых пикселей F_{10} вычисляется в несколько шагов:

Строится гистограмма H_{YGr} значений яркости Y пикселей, принадлежащих зеленом областям;

Вычисляется ширина гистограммы D как расстояние между ее правым и левым ненулевыми элементами;

Признак F_{10} вычисляется как пропорция гистограммы, лежащая не дальше восьмой части ее ширины от ее

$$F_{10} = \frac{\sum_{i=P_8-D/8}^{P_8+D/8} H_{YGr}(i)}{\sum_{i=0}^{255} H_{YGr}(i)}$$

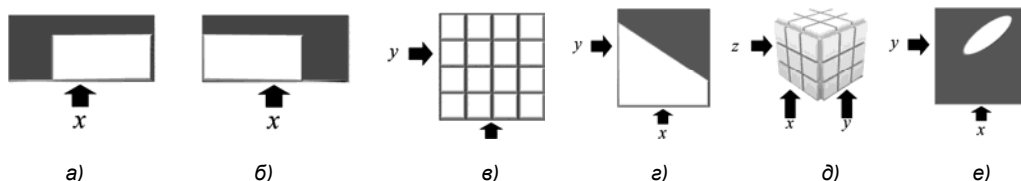


Рис. 3. Элементарные классификаторы

Результирующий классификатор синтезируется в форме ориентированного ациклического графа с элементарными классификаторами в узлах.

Для визуального синтеза классификаторов разработано программное обеспечение, реализующее элементарные классификаторы следующих типов: одномерные пороговые преобразования $\theta_i^{HI}(x)$ и $\theta_i^{LO}(x)$, двумерные пороговые преобразования $\theta_i^{2D}(x, y)$, линейные классификаторы $\theta_i^L(x, y)$, трехмерные пороговые преобразования $\theta_i^{3D}(x, y, z)$ и двумерные эллиптические классификаторы $\theta_i^e(x, y)$. Одномерные пороговые преобразования имеют один входной аргумент и описываются формулой $\theta_i^{HI}(x) = \begin{cases} TRUE, & x > T_i \\ FALSE, & otherwise \end{cases}$ (рис. 3.а)

или $\theta_i^{LO}(x) = \begin{cases} TRUE, & x < T_i \\ FALSE, & otherwise \end{cases}$ (рис. 3.б). Параметры

двумерных пороговых преобразований (или табличных классификаторов) представлены на рис. 3.в. Они описываются двумя векторами порогов: $V_T^1 = [T_0^1 \ T_1^1 \ \dots \ T_N^1]$, где $T_0^1 < T_1^1 < T_2^1 < \dots < T_N^1$, и $V_T^2 = [T_0^2 \ T_1^2 \ \dots \ T_M^2]$, где $T_0^2 < T_1^2 < T_2^2 < \dots < T_M^2$.

Количественные значения порогов для каждого измерения могут не совпадать: $M \neq N$. Выход табличного классификатора имеет форму двумерного вектора размерности $M \times N$: $\theta^{2D}(x, y) = [y_{11}(x, y), y_{12}(x, y), \dots, y_{1N}(x, y), y_{21}(x, y), \dots, y_{M1}(x, y), y_{M2}(x, y), \dots, y_{MN}(x, y)]$, где каждый элемент является логической величиной $y_{ij}(x, y) =$

$= x \geq T_{i-1}^1 \wedge x \leq T_i^1 \wedge y \geq T_{j-1}^2 \wedge y \leq T_j^2$. Выход линейного классификатора вычисляется на основании следующего

выражения: $\theta_i^L(x, y) = K_1 \cdot x + K_2 \cdot y + B > 0$, где K_1, K_2 и B – предопределенные константы (рис. 3.г). Параметры трехмерных табличных классификаторов (рис. 3.д) описываются тремя векторами порогов:

$V_T^1 = [T_0^1 \ T_1^1 \ \dots \ T_N^1]$, где $T_0^1 < T_1^1 < T_2^1 < \dots < T_N^1$;

$V_T^2 = [T_0^2 \ T_1^2 \ \dots \ T_M^2]$, где $T_0^2 < T_1^2 < T_2^2 < \dots < T_M^2$ и

$V_T^3 = [T_0^3 \ T_1^3 \ \dots \ T_L^3]$, где $T_0^3 < T_1^3 < T_2^3 < \dots < T_L^3$.

Выход табличного классификатора имеет вид трехмерного вектора размерности $M \times N \times L$: $\theta^{2D}(x, y) = [y_{111}(x, y, z), y_{112}(x, y, z), \dots, y_{MNL}(x, y, z)]$.

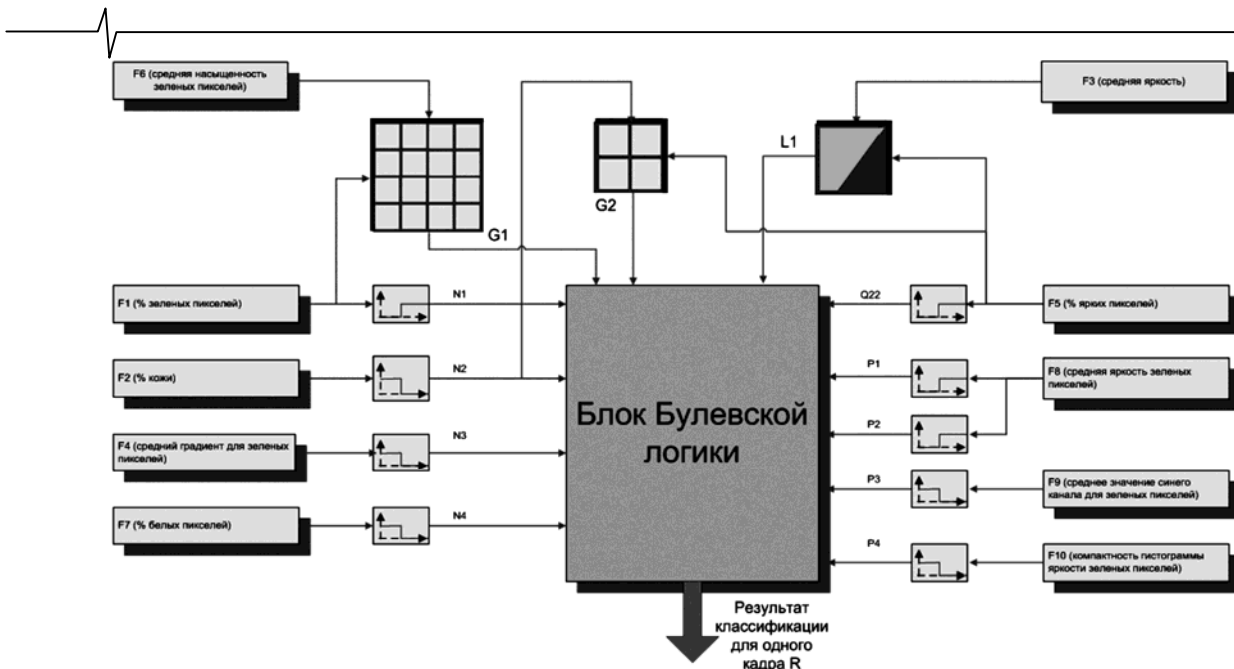


Рис. 4. Структура классификатора для одного кадра

Выход эллиптического классификатора (рис. 3 е) вычисляется как $\theta_i^e(x, y) = \frac{(x' - X_C)^2}{a^2} + \frac{(y' - Y_C)^2}{b^2} < 1$, где $x' = x \cdot \cos \alpha - y \cdot \sin \alpha$, $y' = x \cdot \sin \alpha + y \cdot \cos \alpha$ а X_C, Y_C, a, b и α – предопределенные константы.

В процессе исследований выяснено, что во многих случаях вполне достаточно ограничиться простейшим вариантом классификатора, включающего одномерные и двумерные пороговые преобразования, а также набор линейных классификаторов (рис. 4).

Сформулируем основные правила, на основе которых делается окончательный вывод о том, относится ли текущий кадр к игровым эпизодам. При этом будем считать, что пороговые величины T^1, \dots, T^{13} , $\tilde{T}_i^1, \tilde{T}_i^2, i = 1..4$ и $\hat{T}_i^1, \hat{T}_i^2, i = 1..2$ являются действительными числами, определенными на интервале $[0, 1]$.

Первое правило реализовано в виде табличного классификатора на основе признаков F_1 и F_6 и ориентировано на вычисление вектора

$$G_1 = |y_{11}, y_{12}, y_{13}, y_{14}, y_{21}, y_{22}, y_{23}, y_{24}, y_{31}, y_{32}, y_{33}, y_{34}, y_{41}, y_{42}, y_{43}, y_{44}|.$$

Цель этого правила состоит в том, чтобы выделить случаи дальнего, общего и крупного планов, которым соответствуют ненулевые значения y_{ij} (очевидно, что нулевая норма вектора G_1 равна единице). Результат выполнения этого правила используется в дальнейшем для выбора порогов. При этом если насыщенность зеленого мала, то имеет смысл применить более строгие правила, нежели в случае, когда она высока или принимает средние значения. Второе правило, реализуемое в виде табличного классификатора на основе признаков F_2 и F_5 , ориентировано на определение вектора $G_2 = |z_{11}, z_{12}, z_{21}, z_{22}|$. Применение этого правила обеспечивает в дальнейшем выделение кадров с низкой до-

лей зеленых пикселей.

Случай с очень низкой долей зеленых пикселей рассмотрен отдельно и реализован в виде «негативного» правила: $N_1 = F_1 > T^3$. Случай с очень высокой долей пикселей цвета кожи также контролируется «негативным» правилом: $N_2 = F_2 < T^4$. Текстурные свойства зеленых пикселей контролируются еще одним «негативным» правилом $N_3 = F_4 < T^5$. Игровые эпизоды содержат также определенную (впрочем, сравнительно небольшую) долю белых пикселей: $N_4 = F_7 > T^6$.

Поскольку вспомогательный детектор ярких и насыщенных пикселей основан на фиксированных порогах, очевидно, что результат детектирования будет зависеть от общей яркости изображения.

После применения простой гамма-коррекции изображения доля детектированных ярких и насыщенных пикселей может измениться в разы. Для решения этой проблемы используется линейный классификатор $Q_1 = K_1 \cdot F_3 + K_2 \cdot F_5 + B > 0$, где K_1, K_2 и B – предопределенные константы.

Количество ярких и насыщенных пикселей контролируется правилом $Q_2 = F_5 > T^9$. Средняя яркость зеленых пикселей разбита на два диапазона, для каждого из которых в дальнейшем используется различная эмпирическая логика, например, $P_1 = F_8 > T^{10}$ и $P_2 = F_8 > T^{11}$. Среднее значение синей компоненты для зеленых пикселей контролируется правилом $P_3 = F_9 < T^{12}$, а степень компактности гистограммы яркости зеленых пикселей – $P_4 = F_{10} < T^{13}$. Окончательный результат детектирования R вычисляется по формуле $R = (\neg V_1) \wedge P_1 \wedge P_4 \wedge V_2 \wedge (P_2 \vee (\neg P_2 \wedge P_3))$, где $V_1 = N_1 \vee N_2 \vee N_3 \vee N_4 \vee z_{11}$ и $V_2 = (y_{22} \wedge Q_2) \vee y_{23} \vee (y_{24} \wedge Q_1) \vee y_{32} \vee y_{33} \vee y_{34} \vee y_{42} \vee y_{43} \vee y_{44}$.

Для обеспечения согласованных результатов детек-

тирования в пределах одной сцены, результат детектирования, полученный для первого кадра, распространяется на всю сцену (рис. 5).

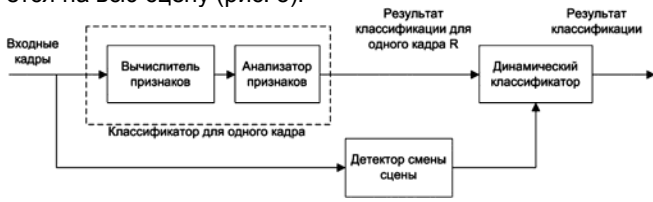


Рис. 5. Структура классификатора для видеопоследовательности

Для детектирования момента смены сцены предложен следующий алгоритм (рис. 6) Значения цветочных каналов RGB текущего видеокadra и центры кластеров K_C с выхода блока задержки подаются на вход блока кластеризации. Центры кластеров представляют собой матрицу размером $K \times 3$, где K – количество класте-

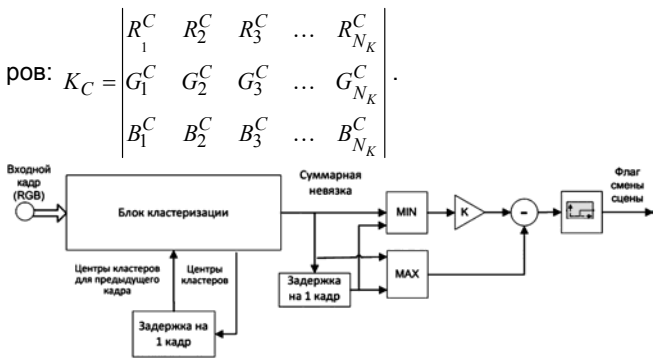


Рис. 6. Структура детектора смены сцены

В настоящем исследовании K принималось равным 8. Кластеризация представляет собой один итерационный шаг метода K средних, когда каждый пиксель $P(i, j) = |R(i, j) G(i, j) B(i, j)|$ с координатами i, j относят к кластеру $K(i, j) = \arg \min_{k=1..N_k} D(P(i, j), C_k)$, где

$C_k = |R_k^C \ G_k^C \ B_k^C|$ – центр k -го кластера и $D(x, y) = \|x - y\|$ – какая-нибудь векторная норма. Суммарная невязка вычисляется при этом как $E = \sum_{i=1..w, j=1..h} D(P(i, j), C_{K(i, j)})$. Центры кластеров обновля-

ются по формуле $\tilde{K}_C = \begin{bmatrix} \tilde{R}_1^C & \tilde{R}_2^C & \tilde{R}_3^C & \dots & \tilde{R}_{N_k}^C \\ \tilde{G}_1^C & \tilde{G}_2^C & \tilde{G}_3^C & \dots & \tilde{G}_{N_k}^C \\ \tilde{B}_1^C & \tilde{B}_2^C & \tilde{B}_3^C & \dots & \tilde{B}_{N_k}^C \end{bmatrix}$, где $\tilde{R}_k^C = \frac{\sum_{i=1..w, j=1..h} \delta(K(i, j)=k) \cdot R(i, j)}{\sum_{i=1..w, j=1..h} \delta(K(i, j)=k)}$, $\tilde{G}_k^C = \frac{\sum_{i=1..w, j=1..h} \delta(K(i, j)=k) \cdot G(i, j)}{\sum_{i=1..w, j=1..h} \delta(K(i, j)=k)}$, $\tilde{B}_k^C = \frac{\sum_{i=1..w, j=1..h} \delta(K(i, j)=k) \cdot B(i, j)}{\sum_{i=1..w, j=1..h} \delta(K(i, j)=k)}$. Флаг смены сцены вы-

числяется по формуле $\frac{\max(E_{cur}, E_{prev})}{\min(E_{cur}, E_{prev})} > T_{break}$, где

E_{cur} – суммарная невязка в текущем кадре, E_{prev} – суммарная невязка в предыдущем кадре а $T_{break} = 8$ – заданный порог. Такой детектор работает более надежно, чем детектор, основанный на гистограмме, и при этом является более дешевым с точки зрения его аппаратной реализации.

Результаты

Для демонстрации работы алгоритма реализована компьютерная программа на языках C++/C#.NET. Значения признаков $F_1, F_2, F_3, F_4, F_5, F_6$ выводятся в программном интерфейсе в форме отдельных графиков, а признаки $F_1, F_2, F_3, F_4, F_5, F_7$ в форме столбиковых эпюр (рис. 7). Классификатор протестирован не менее чем на 20 часах видео в реальном масштабе времени (табл. 2). Ошибки первого рода оценивались путем усреднения выхода классификатора на видеопоследовательностях класса С4. В процессе классификации достигнут 95%-ый порог точности. Для оценки ошибок второго рода выбирались случайные кадры общим числом $N_{total} = 220$ из различных последовательностей классов С1-С3 и точность классификации вычислялась как $\frac{N_{C1}^+ + N_{C2}^+}{N_{total}} \cdot 100\%$, где N_{C1}^+ число кадров класса С1,

классифицированных как «спорт» и N_{C2}^+ число кадров класса С2, классифицированных как «спорт». Точность классификации составила 96.5%. Вычисления с использованием коэффициентов из табл. 1 для кадров класса С3 показали приемлемую результирующую точность (выше 95%), однако эти измерения не имеют особой ценности, поскольку мнения экспертов расходятся по поводу того, как классифицировать многие из изображений класса С3.

Производительность алгоритма на компьютере с 2GHz процессором и 2Mb памяти достигала 30 кадров в секунду. Описанный алгоритм может быть реализован с использованием только сдвигов и сложений, что делает его привлекательным для аппаратной реализации.

Обработка сложных случаев

Несмотря на высокое качество классификации, продемонстрированное выше, требуется дальнейшая работа по устранению замеченных погрешностей классификации (рис. 8). Так, на рис. 8.а фрагмент из фильма о природе был принят за дальний план игры в футбол, а гусеница на рис. 8.б ошибочно принята за крупный план игрока. В следующей версии алгоритма таких ошибок можно будет избежать за счет применения усовершенствованного детектора кожи, добавления детектора белой разметки для дальних планов и добавления кросс-доменного признака, сочетающего текстуру и цвет для области ярких и насыщенных цветов. Ошибка классификации на рис. 8.в вызвана слишком широким толкованием понятия «зеленого» цвета. Для решения этой проблемы целесообразно применить адаптацию детекторов

#	классы	Название	Продолжительность, мин.	Разрешение
1	C1,C3	Футбол: 2006 FIFA World Cup Semifinal: Italy vs. Germany	147	960x544
2	C1,C3	Футбол: Liverpool- Manchester United, BBC broadcast	83	592x336
3	C1,C3	Футбол: Milan – Liverpool	51	640x352
4	C1,C3	Футбол: Manchester United – Villareal (Sky Sports2)	72	624x336
5	C2,C3	Амер. футбол: Penn State - Ohio State	141	352x288
6	C2,C3	Амер. футбол: Sugarbowl	168	352x288
7	C4	Wild South America	100	720x576
8	C4	Greatest places	24	1280x720
9	C4	«Люди в черном»	98	720x480
10	C4	Файлы из youtube, всего 84	421	720x528
		Всего	1305	

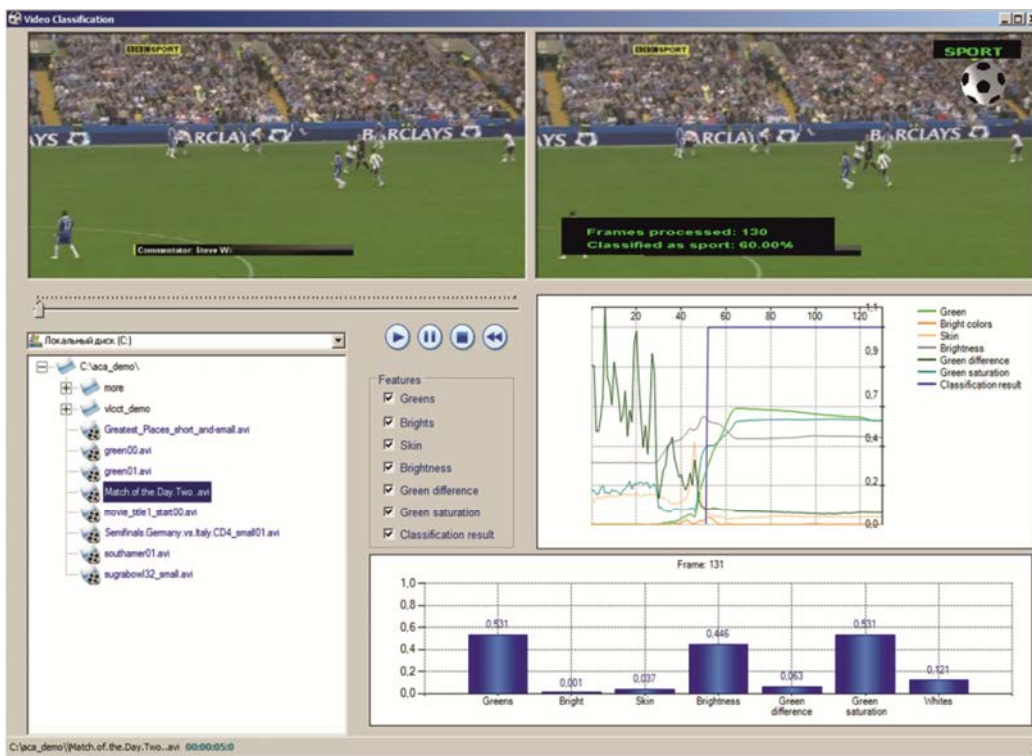
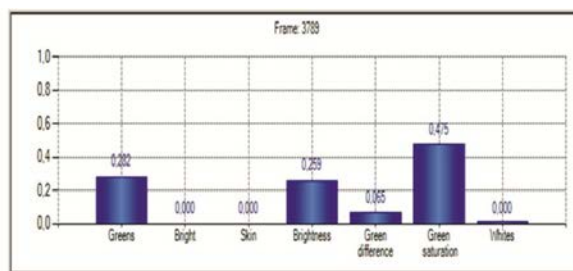
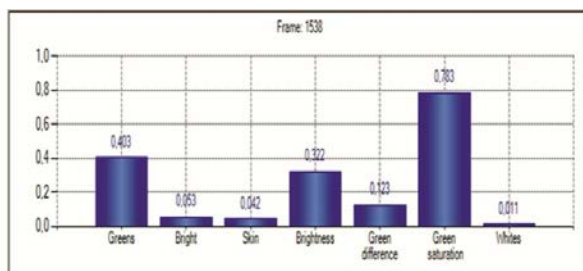


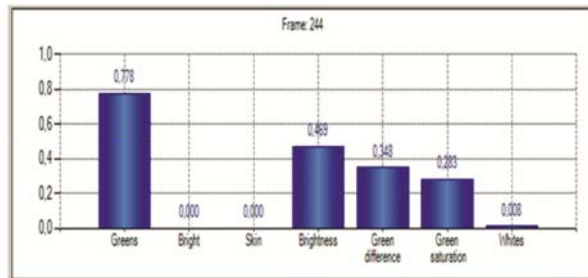
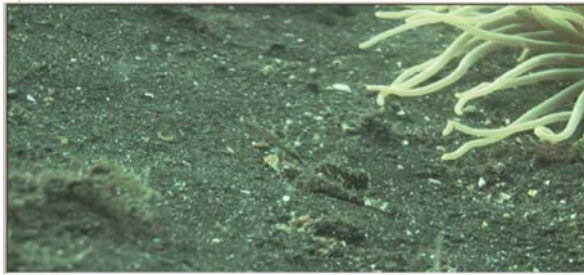
Рис.7. Интерфейс демонстрационной программы



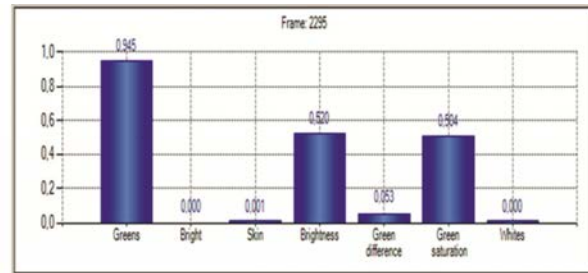
а)



б)



е)



е)

Рис. 8. Ошибки детектирования

цвета к глобальному освещению сцены. Для исправления ошибки, показанной на рис. 8.г придется разработать классификатор силуэтов, что представляет собой значительно более сложную задачу, для которой, скорее всего, не удастся найти приемлемое по производительности решение.

Заключение

Разработан быстродействующий интуитивно понятный алгоритм выделения игровых эпизодов в видеопотоке. Алгоритм дает результат в реальном времени по одному кадру, при этом гарантируется его полноценная реализация на стороне телевизионного приемника в виде блока микросхемы видеобработки.

Литература

1. Brezeale D., Cook D.J., Using Closed Captions and Visual Features to Classify Movies by Genre, Poster session of the Seventh International Workshop on Multimedia Data Mining (MDM/KDD2006), 2006.
2. Liang Bai; Song-Yang Lao; Hu-Xiong Liao; Jian-Yun Chen, Audio Classification and Segmentation for Sports Video Structure Extraction using Support Vector Machine. International Conference on Machine Learning and Cybernetics, Aug. 2006 pp. 3303-3307
3. Dinh P.Q., Dorai C., Venkatesh S. Video genre categorization using audio wavelet coefficients. In 5th Asian Conference on Computer Vision, Melbourne, Australia, Jan 23-25 2002.
4. Roach M., Mason J. Classification of video genre using audio. Eurospeech, 4:2693-2696, 2001.
5. Subashini K., Palanivel S., Ramalingam V., Audio-Video based Classification using SVM and AANN. International Journal of Computer Applications Volume 44- No.6, April 2012. Pp. 33-39
6. Jiang X., Sun T., Chen B., A Novel Video Content Classification Algorithm Based on Combined Visual Features Model. 2nd International Congress on Image and Sig-

nal Processing, 2009. CISP '09. 17-19 Oct. 2009, pp. 1-6.

7. Huang H.Y., Shih W.S., Hsu W.H., A Film Classifier Based on Low-level Visual Features, Journal of Multimedia, vol. 3, no. 3, July 2008

8. Yuan Y., Wan C., The application of edge feature in automatic sports genre classification. IEEE Conference on Cybernetics and Intelligent Systems, 2004, Vol. 2, pp. 1133-1136.

9. Wei G., Agnihotri L., Dimitrova N., Tv Program Classification Based On Face And Text Processing. IEEE International Conference on Multimedia and Expo, 2000. ICME 2000. Vol. 3, pp. 1345-1348.

10. Y. Liu; J.R. Kender, Video frame categorization using sort-merge feature selection Proceedings. Workshop on Motion and Video Computing, 2002. Volume, Issue, 5-6 Dec. 2002 pp: 72 – 77.

11. Takagi S., Hattori S.M., Yokoyama, K.; Kodate, A.; Tominaga, Sports video categorizing method using camera motion parameters H. International Conference on Multimedia and Expo, 2003. ICME apos;03. Proceedings. 2003 Volume 2, Issue , 6-9 July 2003 Page(s): II - 461-4 vol.2

12. Statistical analyzing method of camera motion parameters for categorizing sports video Takagi, S.; Hattori, S.; Yokoyama, K.; Kodate, A.; Tominaga, H. International Conference on Visual Information Engineering, 2003. VIE 2003. 7-9 July 2003 pp. 222-225

13. Gillespie, W.J.; Nguyen, D.T., Classification of video shots using activity power flow First IEEE Consumer Communications and Networking Conference, 2004. Volume 5-8 Jan. 2004, Page(s): 336-340

14. Kittler J., Messer K., Christmas W., Levinaise-Obada, Kourbaroulis D. Generation of Semantic Cues for Sports Video Annotation. In ICIP, pp. 26-29, 2001. 7new11. Jaser, E. Kittler, J. Christmas, W., Hierarchical decision making scheme for sports video categorisation with temporal post-processing IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004. 27 June-2 July 2004, Vol. 2, pp. II-908- 913



15. Jaser, E. Kittler, J. Christmas, W., Hierarchical decision making scheme for sports video categorisation with temporal post-processing IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004. 27 June-2 July 2004, Vol. 2, pp. II-908-913
16. Choroś K., Pawlaczyk P., Content-Based Scene Detection and Analysis Method for Automatic Classification of TV Sports News. Rough Sets and Current Trends in Computing Lecture Notes in Computer Science Volume 6086, 2010, pp. 120-129
17. Pass G., Zabih R., Miller J. Comparing images using color coherence vectors, In Proc. Proceedings of the fourth ACM international conference on Multimedia (1996)
18. Ionescu B. E., Rasche C., Vertan C., Lambert P., A Contour-Color-Action Approach to Automatic Classification of Several Common Video Genres. Adaptive Multimedia Retrieval. Context, Exploration, and Fusion. Lecture Notes in Computer Science Volume 6817, 2012, pp 74-88
19. Vaswani N., Chellappa R., Principal Components Null Space Analysis for Image and Video Classification IEEE Transactions on Image Processing, Vol. 15, no. 7, July 2006, pp. 1816-1830
20. Koskela M., Sjöberg M., Laaksonen J. Improving Automatic Video Retrieval with Semantic Concept Detection. Lecture Notes in Computer Science Volume 5575, 2009, – pp 480-489
21. Truong B.T., Venkatesh S., Dprai C., Automatic Genre Identification for Content-Based Video Categorization. 15th International Conference on Pattern Recognition (ICPR'00) – Volume 4, p. 4230. 22. SEEMORE: combining color, shape, and texture histogramming in a neurally inspired approach to visual object recognition. (1997). Neural Computation, 9(4), 777-804. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9161022>
23. Machajdik J., Hanbury A., Affective Image Classification using Features Inspired by Psychology and Art Theory Proceedings of the international conference on Multimedia, ACM New York, NY, USA 2010. pp. 83-92
24. Li-Jia Li, Hao Su, Eric. P. Xing, Li Fei-Fei. Object Bank: A High-Level Image Representation for Scene Classification and Semantic Feature Sparsification. Proceedings of the Neural Information Processing Systems (NIPS). 2010.
25. Godbole S. Exploiting confusion matrices for automatic generation of topic hierarchies and scaling up multi-way classifiers. Indian Institute of Technology – Bombay. Annual Progress Report. January 2002. <http://www.it.iitb.ac.in/~shantanu/work/aps2002.pdf>
26. Gomez G., Sanchez M., Sucar L.E., On selecting an appropriate color space for skin detection, Springer-Verlag: Lecture Notes in Artificial Intelligence, vol. 2313, 2002, pp. 70–79.

У в а ж а е м ы е а в т о р ы !

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

1) 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 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.

3) Требования к иллюстрациям:

- векторные (схемы, графики) – желательно использование графических редакторов Adobe Illustrator или Corel DRAW;
- растровые (фотографии, рисунки) – М 1:1, разрешение не менее 300 dpi, формат tiff.

УДК:004.932.4

МЕТОДИКА ОБЪЕКТИВНОГО СРАВНЕНИЯ АЛГОРИТМОВ МАТИРОВАНИЯ ВИДЕО

Ерофеев М.В., аспирант МГУ им. М. В. Ломоносова, факультет вычислительной математики и кибернетики, лаборатория компьютерной графики, e-mail: merofeev@graphics.cs.msu.ru;

Гитман Ю.А., студент МГУ им. М. В. Ломоносова, e-mail: ygitman@graphics.cs.msu.ru;

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

Федоров А.А., аспирант МГУ им. М. В. Ломоносова, e-mail: afedorov@graphics.cs.msu.ru.

OBJECTIVE VIDEO MATTING METHODS' COMPARISON TECHNIQUE

Erofeev M., Gitman Y., Vatolin D., Fedorov A.

Formally, matting is a problem of decomposition of image into foreground image, background image and foreground transparency map. Until now there was only common method of image matting comparison not applicable to video matting comparison. Moreover, authors of video matting methods either do not perform any objective evaluation or compare their method to one or two competitors.

In this paper we propose video matting methods' comparison technique by spatial error and temporal coherence. To get ground-truth for our test data set we employed chroma keying and designed stop-motion capture procedure. Finally, we show comparison results of 12 matting methods.

Keywords: video matting, trimap, objective quality estimation.

Ключевые слова: матирование видео, тернарная маска, объективная оценка качества.

Введение

Одной из ключевых задач, возникающих при редактировании и монтаже изображений и видеопоследовательностей, является построение карты прозрачности (матирования) объекта переднего плана для последующей замены фона, элементов фона или изменения положения объекта относительно фона.

Формальная постановка задачи матирования имеет следующий вид: разделить данное изображение I (рис. 1а) на изображение объекта переднего плана F (рис. 1г), изображение заднего плана B и карту прозрачности переднего плана α (рис. 1в) таким образом, чтобы было верно следующее уравнение:

$$I = \alpha F + (1 - \alpha)B. \quad (1)$$

Очевидно, что в таком виде задача является некорректно поставленной и имеет множество решений, включая тривиальное ($I=F$, $\alpha=1$). С целью найти единственное решение, имеющее высокое визуальное качество с точки зрения пользователей, авторы работ в данной области делают дополнительные предположения о

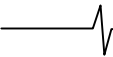
Задача матирования состоит в разделении данного изображения или видео на передний план, задний план и карту прозрачности переднего плана. До сегодняшнего дня существовала только общепринятая методика объективного сравнения алгоритмов матирования изображений, неприменимая для сравнения методов матирования видео. Авторы алгоритмов матирования видео сообщали лишь результаты сравнения их метода с одним или двумя аналогами на малом тестовом наборе или вообще не сообщали таковых.

Данная работа предлагает методику сравнения алгоритмов матирования по их пространственной ошибке и стабильности результата во времени. А также содержит результаты проведенного сравнения ряда современных алгоритмов матирования видео и изображений. Для создания тестовой выборки была использована цветная электронная риппроекция, а также оригинальная методика на основе кукольной мультипликации.

непрерывности результатов [1], близких к нулю модулям градиентов F и B [2], а также используют дополнительные входные данные, например, тернарную маску [3]. Тернарная маска (рис. 1б) представляет собой карту, содержащую три уровня: гарантированный передний план (отмечен белым цветом), гарантированный задний план (отмечен черным цветом) и неизвестная область (отмечена серым цветом), для которой требуется решить задачу матирования. Пример тернарной маски изображен на рис. 1б.



Рис. 1. Пример результата работы алгоритма матирования: а-б – входные данные, в-г – выходные данные



Задача матирования видео является обобщением задачи матирования изображений на видеопоследовательности. Такое обобщение порождает дополнительные требования к алгоритмам матирования видео:

1. Результирующая карта прозрачности должна быть стабильна во времени, ввиду высокой чувствительности зрительной системы человека к различиям между соседними кадрами.

2. Алгоритм не должен требовать значительных усилий со стороны пользователя. Например, недопустимо требовать тернарную маску для каждого кадра обрабатываемой видеопоследовательности.

3. Алгоритм должен иметь адекватную объемам видеоданных вычислительную сложность.

Упомянутые выше требования сдерживали развитие алгоритмов матирования видео в сравнении с алгоритмами матирования изображений. Тем не менее, в последние годы имел место ряд достижений и в области матирования видео [4–8].

Стоит упомянуть, что появление открытого для новых участников сравнения алгоритмов матирования изображений [9] привело к существенному прогрессу в этой области. Тем не менее, ввиду дополнительных требований, предъявляемых к алгоритмам матирования видео, это сравнение не применимо для их оценки.

Данная статья предлагает решение двух задач, возникающих при проведении сравнения алгоритмов матирования: подготовка эталонных данных и выбор метода оценки качества.

Для создания эталонных карт прозрачности, необходимых для проведения объективного сравнения, было использовано две методики:

1. Цветная электронная риппроекция (chromakeying), широко используемая в киноиндустрии.
2. Кукольная мультипликация.

Обзор области исследования

Методы сравнения алгоритмов матирования

Открытое сравнение алгоритмов матирования изображений, предложенное в [9], оказало значительное влияние на данное исследование. Авторы этого сравнения предложили методику получения эталонных данных карт прозрачности, подготовили набор сложных тестовых изображений, предложили метрики качества, отвечающие особенностям субъективного восприятия изображений человеком в ряде частных случаев. Предложенное сравнение стало общепризнанным методом оценки качества алгоритмов матирования изображений.

Несмотря на отсутствие аналогичного сравнения алгоритмов матирования видео, некоторые авторы в своих работах проводят самостоятельную объективную оценку предлагаемого ими метода. Методики оценки, использованные в данных работах, могут быть классифицированы по следующим критериям:

Способ получения эталонных карт прозрачности

В работах [6, 7] тестовая выборка была сформирована путем наложения видеопоследовательностей с картой прозрачности, полученной с помощью электронной риппроекции, на новый фон. Авторы [10] использовали компьютерную графику для построения

синтетических видеопоследовательностей с известной картой прозрачности переднего плана.

Метрика пространственной ошибки

Для оценки пространственной ошибки исследуемых методов в работе [7] используется L_1 расстояние между построенной и эталонной картами прозрачности, а в [6, 10] – L_2 расстояние.

Метрика временной стабильности

Для оценки стабильности результата во времени в работе [6] использовано L_2 расстояние между производной по времени вычисленной и эталонной карты прозрачности. Производная по времени пространственной ошибки была применена в [7]. Авторы работ [10, 11], ввиду отсутствия в их распоряжении эталонных карт прозрачности, для оценки временной стабильности вычисляли отношение модуля производной по времени вычисленной карты прозрачности к модулю производной по времени входного изображения.

Алгоритмы, включенные в сравнение

Лишь в работе [6] предлагаемый метод сравнивается с конкурирующими подходами матирования видео. Авторы работ [7, 10, 11] проводят сравнение только с предлагаемым алгоритмом, примененным к каждому кадру независимо.

Стоит также отметить, что большинство статей, предлагающих алгоритмы матирования видео, не содержат результатов объективной оценки качества [4, 5, 8, 12–15].

Целью данной статьи является введение общепризнанного метода объективной оценки качества алгоритмов матирования видео.

Краткий обзор алгоритмов матирования видео

Методы, используемые в существующих алгоритмах матирования видео, призваны обеспечить стабильность получаемой карты прозрачности во времени, а также сократить объем дополнительных входных данных, требуемых от пользователя. Эти методы можно разделить на методы, явно требующие стабильность карты прозрачности во времени, и методы, требующие стабильность внутренних параметров алгоритма, используемых для получения карты прозрачности, то есть требующие стабильность карты прозрачности неявно.

Явные методы, используемые в алгоритмах матирования видео для обеспечения временной устойчивости:

Пространственно-временной матирующий лапласиан

Непосредственное обобщение матирующего лапласиана, первоначально предложенного в [1], было применено в [11], где были сформулированы требования на гладкость карты прозрачности для пространственно-временных окон фиксированной формы. Авторы работ [5–7, 10] применили пространственно-временные окна, форма которых определялась оптическим потоком, что позволило им добиться лучших результатов на видео с существенным движением матируемого объекта. Пространственно-временной нелокальный матирующий лапласиан был

предложен в [8], его авторы заменили требование локальной гладкости карты прозрачности требованием схожести каждого пиксела карты прозрачности с его k ближайшими соседями в пространстве текстур.

Темпоральная составляющая в оптимизируемой энергии

В работах [6,15] в функцию энергии, предложенную в [1], было добавлено дополнительное слагаемое, требующее близости карты прозрачности текущего кадра к карте прозрачности предыдущего кадра, деформированной согласно векторам оптического потока.

Темпоральная пост-фильтрация карты прозрачности

Авторы работы [4] использовали дистантное преобразование, применяемое независимо к каждой линии уровня карты прозрачности, для повышения временной стабильности результатов работы алгоритма матирования изображений. Эта методика впоследствии была использована в профессиональной программной среде для обработки видео Adobe After Effects;

Неявные методы обеспечения стабильности результата во времени:

В работах [4, 10, 12] вектора оптического потока были использованы для распространения тернарной маски ключевых кадров, размеченных пользователем, на всю видеопоследовательность.

Авторы работ [6, 7, 15] вместо независимого построения модели цветов переднего и заднего плана для локальных участков обрабатываемого кадра использовали пространственно-временную модель цвета, параметры которой сохраняются между соседними кадрами.

Набор тестовых видео

Карта, полученная с использованием кукольной мультипликации (рис. 2), имеет существенно большую детализацию.

Проведение объективного сравнения алгоритмов матирования требует наличия набора разнообразных тестовых видео, схожих по свойствам с видео, встречающимися в повседневной жизни. Для проведения количественных оценок качества алгоритмов также потребуются эталонные карты прозрачности для тестовых видео.

Для получения эталонных карт прозрачности были применены следующие методики:

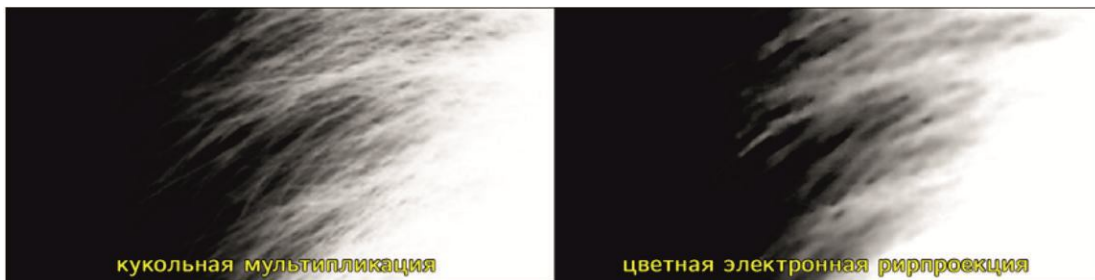


Рис. 2. Карты прозрачности одного и того же фрагмента объекта переднего плана, полученные с использованием методов цветной электронной риппроекции и кукольной мультипликации

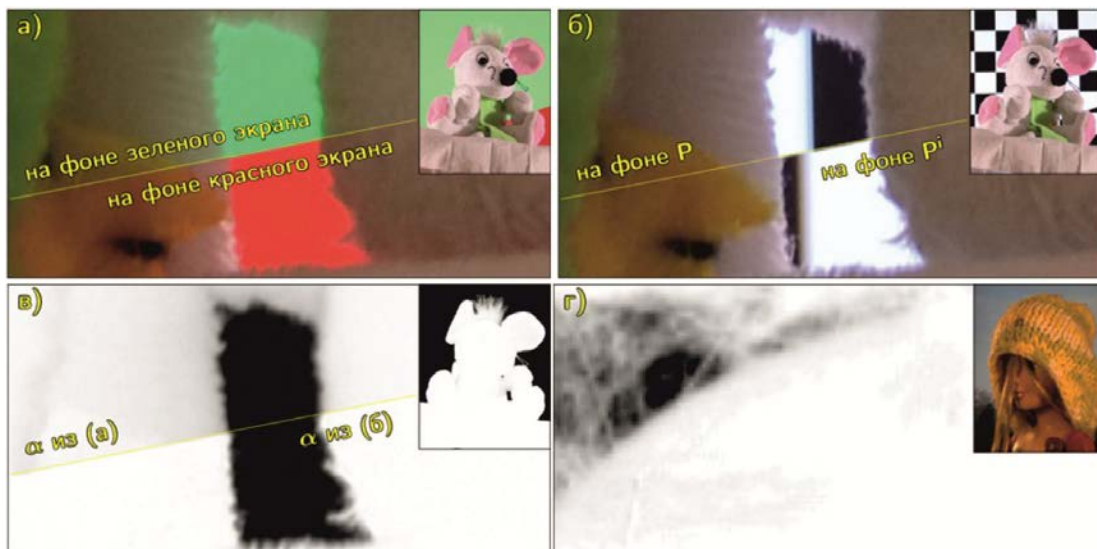


Рис. 3. Проблема влияния монитора на освещение объекта и предлагаемое решение.

- а) Плюшевая игрушка отражает свет от монитора. Верхняя половина изображения была получена при фотографировании объекта на зеленом фоне, нижняя – на красном. б) - Шахматная доска и ее инверсия одинаково освещают игрушку. Верхняя половина изображения была получена при фотографировании объекта на фоне шахматной доски, нижняя – ее инверсии. в) Карты прозрачности, вычисленные из пары изображений (а) (над чертой) и пары изображений (б) (под чертой). Использование шахматной доски и ее инверсии позволило устранить эффект ложной прозрачности. г) Пример проблемы отражений, приводящей к ложной прозрачности на эталонной карте прозрачности из [9]

Цветная электронная рирпроекция [16] позволила получить карты прозрачности для объектов произвольной формы, совершающих движения произвольной сложности.

Кукольная мультипликация позволила получить эталонные карты прозрачности более высокого качества, чем цветная электронная рирпроекция (рис. 2), для ограниченного набора объектов и движений.

Пять последовательностей с эталонной прозрачностью были получены с использованием цветной электронной рирпроекции, для получения других пяти последовательностей была применена кукольная мультипликация. Затем эти последовательности были наложены поверх видеопоследовательностей заднего плана с использованием эталонных карт прозрачности.

Тернарные маски были созданы путем применения операций отсечения по порогу и математической морфологии к эталонным картам прозрачности. Для изучения влияния ширины тернарной маски на качество работы алгоритмов матирования были подготовлены три типа тернарных карт, отличающихся шириной неизвестной области.

Цветная электронная рирпроекция

В данном исследовании были использованы пять последовательностей, снятых на фоне зеленого экрана, содержащих объекты с существенной долей полупрозрачных элементов (волосы, дым, быстро движущиеся объекты). Данные видео были предоставлены компанией Holly wood Camera Work [17]. Объекты переднего плана с соответствующими им картами прозрачности были получены с помощью программного инструмента The Foundry Key light [18]. На некоторых из взятых видеопоследовательностей зеленый экран был неравномерно освещен, в этих случаях для обеспечения высокого качества результата процедура рирпроекции применялась не ко всему кадру целиком, а независимо к небольшим его фрагментам.

Рирпроекция позволила получить карты прозрачности для широкого набора объектов реального мира. Тем не менее, эта процедура не может гарантировать, что полученная карта прозрачности близка к истинной карте прозрачности, так как данный метод основан на предположении, что цвет экрана не содержится в объекте переднего плана. Чтобы получить карты прозрачности, наиболее близкие к истинным, была использована техника кукольной мультипликации.

Кукольная мультипликация

Для получения карт прозрачности высокого качества была разработана следующая процедура: объект с полупрозрачными элементами помещался на специальную платформу перед жидкокристаллическим монитором. Положение платформы изменялось малыми дискретными шагами с помощью сервоприводов, подключенных к компьютеру. После каждого изменения положения платформы цифровая камера, расположенная напротив экспериментальной установки, производила серию снимков неподвижного объекта, после каждого снимка изображение на мониторе позади объекта сменялось на

следующее изображение из заранее подготовленного набора. В конце процедуры объект убирался с платформы, и камера производила фотографирование монитора, поочередно отображающего изображения из набора.

После описанной выше процедуры, согласно [9], значения прозрачности могут быть найдены из следующей системы:

$$\begin{cases} I_1 = \alpha F + (1 - \alpha) B_1, \\ \dots \\ I_n = \alpha F + (1 - \alpha) B_n, \end{cases} \quad (2)$$

где I_1, I_2, \dots, I_n – изображения объекта на фоне различных изображений B_1, B_2, \dots, B_n , αF – это неизвестное изображение объекта переднего плана на абсолютно черном фоне, и α – неизвестная карта прозрачности.

Важно отметить, что уравнения системы 2 содержат неявное предположение, что каждая фотография объекта была сделана в одинаковых условиях освещения. В условиях данного эксперимента это предположение нарушается, так как монитор является относительно сильным источником освещения в пределах данной сцены (рис. 3а), а изображение на нем меняется после каждого снимка. Ввиду вышесказанного, данный подход к поиску α приведет к появлению ложной прозрачности у фрагментов объекта переднего плана, отражающих свет, излучаемый монитором. В частности, такой эффект можно наблюдать на эталонных изображениях, используемых в сравнении алгоритмов матирования изображений [9] (рис. 3г).

Чтобы учесть изменения освещения, добавим в систему 2 новое слагаемое:

$$\begin{cases} I_1 = \alpha F + (1 - \alpha) B_1 + S(B_1 * h), \\ \dots \\ I_n = \alpha F + (1 - \alpha) B_n + S(B_n * h), \end{cases} \quad (3)$$

где S – цветное изображение и $B * h$ – свертка заднего плана с неизвестным линейным фильтром. Каждая точка S зависит от отражающих свойств объекта в этой точке, а также от положения этой точки относительно монитора. $B * h$ зависит от среднего цвета некоторой области изображения B , отображенного на мониторе. Введение нового слагаемого делает систему 3 неопределенной. Поэтому, в отличие от [9], помимо однотонных изображений заднего плана, были также использованы изображения шахматной доски. А именно, производилось последовательное фотографирование объекта напротив черно-белой шахматной доски P , а затем инвертированной шахматной доски P^i . Отметим, что P и P^i имеют почти одинаковый средний цвет, и поэтому верно $P * h \approx P^i * h$ (см. рис. 3б). Таким образом, имеем систему из двух уравнений

$$\begin{cases} I_1 = \alpha F + (1 - \alpha) P + S(P * h), \\ I_2 = \alpha F + (1 - \alpha) P^i + S(P^i * h) \end{cases} \quad (4)$$

Вычтем второе уравнение из первого и выразим α :

$$I_1 - I_2 \approx (1 - \alpha)(P - P^i) \Rightarrow \alpha \approx 1 - \frac{(I_1 - I_2, P - P^i)}{(P - P^i)^2} \quad (5)$$

Данная формула позволяет вычислить α везде, кро-

ме узких полос вдоль границ клеток, где разница между шахматной доской и ее инверсией мала ввиду оптического размытия, что приводит к близости $(P - P^i)^2$ к нулю. Чтобы обойти эту проблему, мы дополнительно снимаем объект на фоне шахматных досок, сдвинутых на полклетки влево, вниз и по диагонали, а также их инверсий. Таким образом, для каждого пиксела получаем хотя бы одну систему уравнений, в которой $(P - P^i)^2$ существенно отличается от нуля. Это позволяет найти значение прозрачности в каждой точке, выбирая систему с наибольшим значением $(P - P^i)^2$.

Тем не менее, вычислив α , невозможно однозначно найти F из системы (4). Наиболее простым способом получить F – найти $\alpha F + S(P * h)$ из системы (4) и считать его примерно равным αF . К сожалению, такой метод приведет к появлению неестественного свечения около границ объекта из-за неоднородности S вблизи границ.

Для устранения этой проблемы производится дополнительное фотографирование объекта на фоне однотонных изображений, что позволяет найти F из системы (3) с использованием ранее вычисленной карты прозрачности α , полагая $B_i * h \approx B_i$, ввиду однородности фона.

Описанная выше процедура позволяет вычислить F , α и S для каждого кадра тестовой видеопоследовательности.

Для подготовки тестов и проведения измерений в рамках предлагаемого сравнения используется только F и α . Тем не менее, S может быть использована для повышения визуального качества изображений переднего плана, помещенных на новый фон путем эмуляции влияния освещения новой сцены на цвет объекта.

Объективное сравнение

В проведенное сравнение вошли как методы матирования видео, так и методы матирования изображений (примененные независимо к каждому кадру). Каждый метод поочередно применялся к тестовым видео из подготовленного ранее набора, в качестве дополнительных данных методу предоставлялась тернарная маска. Все методы, вошедшие в сравнение, требовали тернарную маску для каждого кадра. Полученное изображение переднего плана и карта прозрачности затем сравнивались с эталонными данными с целью объективной оценки качества результата.

Основной областью применения алгоритмов матирования является перенос объекта переднего плана на новый фон. Несмотря на это, в сравнении, проведенном в [9], учитывается лишь близость карты прозрачности к эталону. Это ведет к тому, что качество изображения на новом фоне, которое можно получить с помощью испытываемого метода, оценивается лишь косвенно. Тем не менее, оценить качество изображения на новом фоне для произвольного фона является нетривиальной задачей.

Рассмотрим новый фон (B) как случайную величину с равномерным распределением на отрезке $[0;1]$. Тогда

новое изображение I' , полученное в результате наложения извлеченного переднего плана F' с использованием карты прозрачности α на новый фон, будет функцией этой случайной величины. Пусть также α и F – эталонные карта прозрачности и передний план соответственно. В таком случае математическое ожидание среднеквадратического отклонения полученного изображения на новом фоне от эталонного можно вычислить по формуле:

$$\begin{aligned} E[I - I']^2 &= E[\alpha F + (1 - \alpha)B - \alpha' F' - (1 - \alpha')B]^2 = \\ &= (\alpha F - \alpha' F' + \frac{\alpha' - \alpha}{2})^2 + \frac{(\alpha' - \alpha)^2}{12}. \end{aligned} \quad (6)$$

В рамках данного сравнения значение, вычисленное по приведенной выше формуле и нормированное на число неизвестных пикселей на тернарной маске, используется для оценки пространственной ошибки метода матирования на тестовом видео. В табл. 1 приведены результаты сравнения методов матирования по величине пространственной ошибки.

Важным требованием, предъявляемым к алгоритмам матирования видео, является также стабильность результата во времени. Чтобы оценить стабильность результата во времени, была применена следующая гипотеза: покадровое значение пространственной ошибки у стабильной во времени видеопоследовательности должно оставаться почти постоянным. Тогда оценить стабильность во времени метода можно как дисперсию значений его покадровой ошибки. В табл. 2 приведены результаты сравнения методов по этому параметру. Интересно отметить, что алгоритм матирования видео, примененный в профессиональной среде обработки видео Adobe After Effects [19], и алгоритм матирования изображений, на котором он основан [20], заняли первое и второе место в сравнении по обоим параметрам.

Дополнительно был проведен анализ зависимости значения пространственной ошибки алгоритмов матирования от ширины неизвестной области тернарной маски (см. рис. 4). Стоит отметить, что методы (Bayesian Matting [3], Robust Matting [20], Shared Matting [21]), основанные на сэмплинге цветов переднего и заднего плана из известных областей тернарной маски, менее устойчивы к увеличению ширины неизвестной области, чем методы (Closed Form [1], Learning Based [22], KNN Matting [23], Nonlocal Matting [24]), основанные на интерполяции карты прозрачности в неизвестную область.

Открытое онлайн сравнение

Для упрощения доступа к результатам проведенного объективного сравнения, а также для обеспечения возможности включать в сравнение новые алгоритмы, был подготовлен сайт <http://videomattng.com/>. Сайт содержит графики и таблицы, отражающие качество работы методов согласно различным метрикам качества. Карты прозрачности и объекты переднего плана, вычисленные с использованием методов, вошедших в сравнение, доступны для просмотра с использованием специального видео проигрывателя, поддерживающего одновременный просмотр входных и выходных данных, а также их увеличенных фрагментов.

Таблица 1. Результаты сравнения алгоритмов матирования по величине пространственной ошибки (вычисленной по формуле 6).

Метод	Ср. место	Рирпроекция					Кукольная мультипликация				
		№1	№2	№3	№4	№5	№6	№7	№8	№9	№10
Robust Matting [20]	3.1	0.117	0.121	0.102	0.231	0.155	0.184	0.168	0.154	0.055	0.137
Refine Edge [19]	3.5	0.094	0.142	0.134	0.326	0.099	0.171	0.142	0.133	0.058	0.163
Comprehensive Sampling [25]	4.5	0.138	0.186	0.129	0.351	0.176	0.165	0.227	0.155	0.049	0.089
Shared Matting [21]	5.0	0.133	0.183	0.109	0.244	0.168	0.202	0.203	0.157	0.070	0.127
Bayesian Matting [3]	6.1	0.140	0.193	0.063	0.300	0.129	0.277	0.175	0.233	0.121	0.308
Learning Based [22]	6.3	0.171	0.165	0.243	0.432	0.195	0.163	0.261	0.121	0.059	0.112
Closed Form [1]	6.6	0.173	0.162	0.243	0.432	0.192	0.165	0.252	0.121	0.058	0.114
Sparse coded matting [26]	6.7	0.150	0.220	0.177	0.330	0.187	0.188	0.235	0.169	0.060	0.108
KNN Matting [23]	8.3	0.155	0.394	0.513	0.270	0.161	0.378	0.220	0.254	0.079	0.136
Spectral Matting [27]	8.6	0.212	0.336	0.078	0.409	0.153	0.324	0.228	0.344	0.081	0.396
Nonlocal matting [24]	9.5	0.142	0.343	0.309	0.482	0.158	0.340	0.234	0.232	0.154	0.352
Weighted Color [28]	9.8	0.210	0.286	0.201	0.447	0.341	0.271	0.415	0.266	0.074	0.153

Таблица 2. Результаты сравнения алгоритмов матирования по стабильности результата во времени (дисперсия покадровой величины пространственной ошибки, вычисленной по формуле 6).

Метод	Ср. место	Рирпроекция					Кукольная мультипликация				
		№1	№2	№3	№4	№5	№6	№7	№8	№9	№10
KNN Matting [23]	2.6	0.004	0.025	0.005	0.006	0.004	0.003	0.005	0.002	0.003	0.016
Robust Matting [20]	3.9	0.009	0.014	0.003	0.013	0.004	0.004	0.003	0.004	0.003	0.020
Refine Edge [19]	4.8	0.003	0.020	0.018	0.009	0.004	0.006	0.006	0.003	0.005	0.016
Comprehensive Sampling [25]	5.2	0.011	0.015	0.007	0.015	0.006	0.003	0.006	0.003	0.004	0.018
Learning Based [22]	5.3	0.018	0.013	0.007	0.007	0.007	0.007	0.006	0.003	0.005	0.018
Closed Form [1]	5.5	0.019	0.013	0.008	0.006	0.007	0.008	0.005	0.003	0.005	0.018
Shared Matting [21]	5.5	0.010	0.022	0.003	0.010	0.007	0.005	0.004	0.003	0.006	0.021
Sparse coded matting [26]	6.5	0.011	0.022	0.009	0.014	0.008	0.004	0.006	0.002	0.004	0.020
Bayesian Matting [3]	7.7	0.012	0.031	0.008	0.009	0.004	0.008	0.006	0.006	0.013	0.018
Weighted Color [28]	10.1	0.015	0.021	0.028	0.022	0.021	0.026	0.015	0.005	0.007	0.025
Nonlocal matting [24]	10.2	0.016	0.030	0.065	0.013	0.016	0.012	0.012	0.007	0.012	0.021
Spectral Matting [27]	10.7	0.027	0.099	0.018	0.010	0.024	0.012	0.012	0.032	0.009	0.071

Сайт также содержит инструкции для желающих протестировать разработанный ими алгоритм матирования видео, а также ссылки на тестовые видеопоследовательности.

На каждом видео из тестового набора пространственная ошибка измерялась независимо. В качестве

дополнительных входных данных методам предоставлялась узкая тернарная маска. Среднее место, которое метод занимает в каждом тесте, указано во второй колонке. Жирным шрифтом выделен лучший результат по каждому тесту.

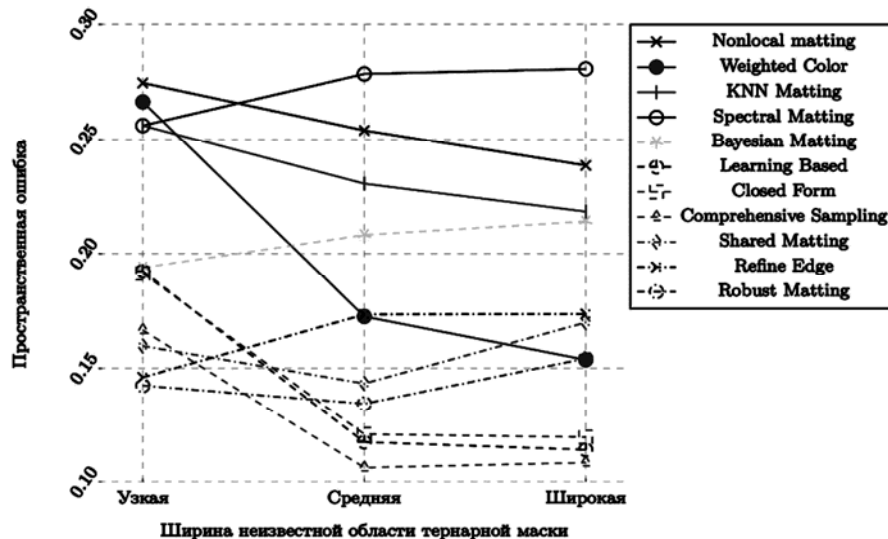


Рис. 4. Влияние ширины области, отмеченной как неизвестная на тернарной маске (рис. 16), на величину пространственной ошибки (вычисленной по формуле 6), нормированной на число неизвестных пикселей

Заключение

В рамках данной работы была предложена оригинальная методика получения эталонных карт прозрачности, превосходящая по качеству методику, примененную в [9]. С использованием этой методики, а также с использованием цветной эталонной рирпроекции, были подготовлены 10 тестовых видео с эталонными картами прозрачности. В статье предложены методы оценки пространственной ошибки и временной стабильности алгоритмов матирования. С помощью данных методов было проведено объективное сравнение существующих алгоритмов матирования видео и изображений. Результаты проведенного сравнения были помещены на интерактивный сайт <http://videomattng.com> для облегчения их анализа.

Основным направлением дальнейшего исследования является проведение экспериментов по субъективной оценке качества методов матирования с использованием репрезентативной выборки респондентов. Результаты таких экспериментов могут быть использованы для верификации предложенных методов объективной оценки, а также для разработки новых методов.

Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта №15-01-08632 а.

Литература

1. Levin A., Lischinski D., Weiss Y. A closed-form solution to natural imagematting // *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*. – 2008. – Vol. 30, no. 2. – P. 228–242.
2. Poisson matting / Jian Sun, Jiaya Jia, Chi-Keung Tang, Heung-Yeung Shum // *ACM Transactions on Graphics (TOG)*. – 2004. – Vol. 23, no. 3. – P. 315–321.
3. A bayesian approach to digital matting / Yung-Yu Chuang, Brian Curless, David H. Salesin, Richard Szeliski // *Computer Vision Pattern Recognition (CVPR)*. – Vol. 2. – 2001. – P. II–264–II–271.
4. Bai X., Wang J., Simons D. Towards temporally-coherent video matting // *International Conference on Computer Vision (ICCV)*. – 2011. – P. 63–74.
5. Sindeev M., Konushin A., Rother C. Alpha-flow for video matting // *Asian Conference on Computer Vision (ACCV)*. — 2013. – P. 438–452.
6. Temporally coherent and spatially accurate video matting / Ehsan Shahrian, Brian Price, Scott Cohen, Deepu Rajan // *Computer Graphics Forum*. – 2014. – Vol. 33, no. 2. – P. 381–390.
7. Video matting via opacity propagation / Zhen Tang, Zhenjiang Miao, Yanli Wan, Dianyong Zhang // *The Visual Computer*. – 2012. – Vol. 28, no. 1. – P. 47–61.
8. Choi I., Lee M., Tai Y.-W. Video matting using multi-frame nonlocal matting laplacian // *European Conference on Computer Vision (ECCV)*. – 2012. – P. 540–553.
9. A perceptually motivated online benchmark for image matting / Christoph Rhemann, Carsten Rother, Jue Wang et al. // *Computer Vision Pattern Recognition (CVPR)*. – 2009. – P. 1826–1833.
10. Lee S.-Y., Yoon J.-C., Lee I.-K. Temporally coherent video matting // *Graphical Models*. – 2010. – Vol. 72, no. 3. – P. 25–33.
11. Spatio-temporally coherent interactive video object segmentation via efficient filtering / Nicole Brosch, Asmaa Hosni, Christoph Rhemann, Margrit Gelautz // *Pattern Recognition*. – Vol. 7476. – 2012. – P. 418–427.
12. Video matting of complex scenes / Yung-Yu Chuang, Aseem Agarwala, Brian Curless et al. // *ACM Transactions on Graphics (TOG)*. – 2002. – Vol. 21, no. 3. – P. 243–248.
13. Apostoloff N., Fitzgibbon A. Bayesian video matting using learnt image priors // *Computer Vision Pattern Recognition (CVPR)*. – Vol. 1. – 2004. – P. I–407–I–414.
14. Corrigan D., Robinson S., Kokaram A. Video matting using motion extended grabcut // *European Conference on Visual Media Production (CVMP)*. – 2008. – P. 3–3(1).
15. Video snapcut: Robust video object cutout using localized classifiers / Xue Bai, Jue Wang, David Simons, Guillermo Sapiro // *ACM Transactions on Graphics (TOG)*. – 2009. – Vol. 28, no. 3. – P. 70:1–70:11.
16. Мамросенко К.А., Гиацинтов А.М. Метод рирпроекции в подсистеме визуализации тренажерно-обучающей системы // *Программные продукты и системы*. 2014. №4.
17. Hollywood camera work. – <http://www.hollywoodcamera.com/greenscreenplates.html>. – Accessed: 2015-01-03.
18. Keylight. – <http://www.thefoundry.co.uk/products/plugins/keylight/>. – Accessed: 2015-01-03.
19. <http://www.adobe.com/en/products/aftereffects.html>.
20. Wang J., Cohen M. F. Optimized color sampling for robust matting // *Computer Vision Pattern Recognition (CVPR)*. – 2007. – P. 1–8.
21. Gastal E. S., Oliveira M. M. Shared sampling for real-time alpha matting // *Computer Graphics Forum*. – 2010. – Vol. 29, no. 2. – P. 575–584.
22. Zheng Y., Kambhampettu C. Learning based digital matting // *International Conference on Computer Vision (ICCV)*. – 2009. – P. 889–896.
23. Chen Q., Li D., Tang C.-K. KNN matting // *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*. – 2013. – Vol. 35, no. 9. – P. 2175–2188.
24. Lee P., Wu Y. Nonlocal matting // *Computer Vision Pattern Recognition (CVPR)*. – 2011. – P. 2193–2200.
25. Improving image matting using comprehensive sampling sets / E. Shahrian, D. Rajan, B. Price, S. Cohen // *Computer Vision Pattern Recognition (CVPR)*. – 2013. – P. 636–643.
26. Johnson J., Rajan D., Cholakkal H. Sparse codes as alpha matte // *British Machine Vision Conference (BMVA)*. – Vol. 32. – 2014. – P. 245–253.
27. Levin A., Rav Acha A., Lischinski D. Spectral matting // *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*. – 2008. – Vol. 30, no. 10. – P. 1699–1712.
28. Shahrian E., Rajan D. Weighted color and texture sample selection for image matting // *Computer Vision Pattern Recognition (CVPR)*. – 2012. – P. 718–725.

ИТЕРАЦИОННЫЙ АЛГОРИТМ ЛОКАЛИЗАЦИИ ЦЕНТРОВ ГЛАЗ НА ИЗОБРАЖЕНИИ ЛИЦА

Хрящев В.В., к.т.н., доцент, доцент кафедры динамики электронных систем Ярославский государственный университет им. П.Г. Демидова, e-mail: vhr@yandex.ru;

Приоров А.Л., д.т.н., доцент, доцент кафедры ДЭС ЯрГУ им. П.Г. Демидова, e-mail: andcat@yandex.ru;

Никитин А.Е., аспирант кафедры ДЭС ЯрГУ им. П.Г. Демидова, e-mail: dcslab@uniyar.ac.ru;

Степанова О.А., аспирант кафедры ДЭС ЯрГУ им. П.Г. Демидова, e-mail: dcslab@uniyar.ac.ru.

ITERATIVE ALGORITHM FOR THE EYE CENTERS LOCALIZATION ON FACIAL IMAGE

Khryashchev V., Priorov A., Nikitin A., Stepanova O.

A new iterative algorithm for the eye centers localization based on multi-block local binary patterns is proposed. This algorithm is based on machine learning using multi-block local binary patterns, previously used in the problem of face detection. The results of the comparative analysis of the given algorithm with other known methods of eyes localization (gradient and Bayesian) in the standard test image database – FERET and BioID are considered. Low computational complexity of considered algorithm can be used in different applications with real-time video processing.

Keywords: face recognition, eye center localization, multi-block local binary pattern, machine learning.

Ключевые слова: распознавание лиц, локализация центров глаз, мультиблочные локальные бинарные шаблоны, машинное обучение.

Введение

Важнейшую роль в современных системах охранного телевидения играют алгоритмы цифровой обработки изображений и технического зрения [1–3], позволяющие контролировать сотни и тысячи видеоканалов в режиме реального времени. Одним из важнейших направлений модернизации таких систем является решение задачи автоматического распознавания объектов [4–6]. Это необходимо для разработки и производства систем, способных интеллектуально оценивать внешнюю среду и выполнять в ней те или иные действия.

Задача точного определения положения глаз на изображении лица, или локализация глаз, важна для широкого спектра современных задач в области компьютерного зрения, таких как определение направления взгляда и угла поворота головы к камере, анализа выражения лица, мимики и т.д. [6–9]. Помимо этого, локализация глаз успешно используется как предварительный этап при распознавании лиц – координаты центров глаз помогают правильно провести нормировку изображения лица после его детектирования (рис. 1) [7, 8]. Исследования показывают, что точность локализации глаз оказывает существенное влияние на качество системы распознавания лиц [10, 11].

Довольно распространено ошибочное мнение, что локализация глаз является довольно простой задачей, так как глаза – всего лишь один из элементов лица. Однако глаза имеют свою уникальную геометрическую структуру, характеристики движения, фотометрические свойства, что делает их довольно трудным объектом

Предлагается итерационный алгоритм локализации центров глаз на основе мультиблочных локальных бинарных шаблонов. Основой для предлагаемого алгоритма является метод на основе машинного обучения с использованием мультиблочных локальных бинарных шаблонов, ранее применявшийся в задаче детектирования лиц. Приводятся результаты сравнительного анализа предложенного алгоритма с другими известными методами локализации (градиентным и байесовским) на стандартных тестовых базах изображений FERET и BioID. Низкая вычислительная сложность рассмотренного алгоритма позволяет использовать его в приложениях обработки видеоданных реального времени.

для исследований. Можно перечислить ряд факторов, которые оказывают значительное влияние на внешний вид глаз на изображении [12–16]:

- большое разнообразие цвета и видов глаз;
- различные эмоции и мимика: к примеру, у смеющегося человека глаза могут быть практически закрыты;
- наличие оптических препятствий на лице: глаза могут быть частично закрыты очками либо прядями волос;
- поза человека: различные варианты взаимного расположения человека и камеры (фронтальное, в профиль, снизу-вверх и т.д.) влияют на внешний вид глаз; например, при расположении головы в профиль один глаз может быть полностью закрыт;

- условия съемки и качество изображения: факторы внешней среды, например, освещение (его спектральный состав, расположение источников освещения и интенсивность) может значительно менять внешний вид глаз. Более того, очень распространенные на практике низкое разрешение изображений, размытость или низкая детализация текстуры могут сильно влиять на качество исходных данных, что становится большой проблемой для любых алгоритмов локализации глаз.

За последние три десятилетия предложено множество различных подходов к решению задачи локализа-

ции глаз. Тем не менее, несмотря на значительные успехи в данной области, стоит отметить, что данная проблема все еще далека от решения. Большинство методов локализации глаз, предложенных исследователями в последнее время, можно условно разделить на три категории [17]:

- методы, основанные на измерении параметров составных элементов глаза;
- методы, создающие статистическую модель глаз на основе обучения;
- методы, использующие информацию о пространственной структуре лица.

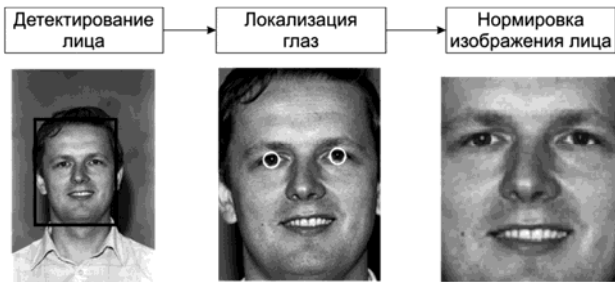


Рис. 1. Роль локализации положения глаз в процедуре распознавания лиц

Проведенный авторами анализ известных алгоритмов локализации глаз показал, что существующие алгоритмы при определении положения центров глаз на изображении лица могут допускать ошибки [11]. Неточности локализации известных алгоритмов связаны, прежде всего, с наличием на лице очков и недостаточным качеством изображений, имеющихся на практике. Погрешности в определении положения глаз серьезно увеличивают число ошибок всей системы распознавания лиц на телевизионном изображении. В работе предлагается новый алгоритм локализации положения центров глаз на основе локальных бинарных паттернов [18, 19], подстраивающийся к качеству и сложности изображения.

Для проведения тестирования использовались два известных алгоритма локализации глаз – градиентный [20] и байесовский [21] – в настоящее время часто используемых в практических приложениях и демонстрирующих приемлемую точность локализации. Градиентный алгоритм локализации использует априорную информацию о пространственной структуре лица. Байесовский алгоритм создан на основе статистического обучения на имеющейся выборке изображений глаз.

Алгоритм локализации центров глаз на основе мультиблочных локальных бинарных шаблонов

Мультиблочный локальный бинарный шаблон (МБ-ЛБШ) является оператором, применяемым к некоторой прямоугольной области изображения (рис. 2). Область разбивается на девять равных блоков. Средняя интенсивность g_c центрального блока рассматривается в качестве порогового значения, которое сравнивается с интенсивностями оставшихся восьми областей окрестности $\{g_1, \dots, g_8\}$ [19].

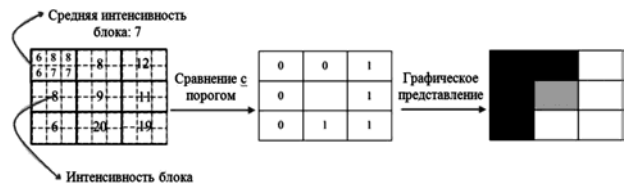


Рис. 2. Пример вычисления оператора МБ-ЛБШ

Выходное значение оператора МБ-ЛБШ центрального блока представляется в следующем виде:

$$МБ-ЛБШ_c = \sum_{i=1}^8 2^i s(g_i - g_c), \tag{1}$$

где s – сигма-функция:

$$s(z) = \begin{cases} 1, & z \geq 0 \\ 0, & z < 0. \end{cases} \tag{2}$$

Представим некоторое изображение в виде вектора признаков $\mathbf{x} = (x^1, \dots, x^S)$. Значения признаков x^1, \dots, x^S являются значениями всех возможных МБ-ЛБШ. Общее число возможных мультиблочных ЛБШ S определяется размером изображений. Так, для изображений разрешением 8x6 пикселей число возможных МБ-ЛБШ равно 45 (24 блока 3x3, 12 блоков 6x3, 6 блоков 3x6 и 3 блока 6x6, как показано на рис. 3). Число возможных признаков быстро возрастает с увеличением разрешения изображений (так, для изображений разрешением 21x15 это число $S = 2450$, а изображений 21x21 – $S = 4900$), поэтому при обучении лучше использовать изображения небольшого разрешения.

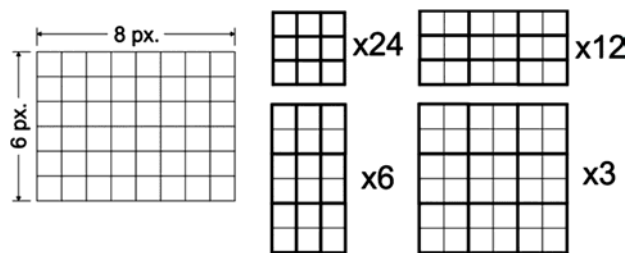


Рис. 3. Возможные МБ-ЛБШ для изображения разрешением 8x6 пикселей

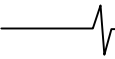
Построение и обучение классификатора

Построим классификатор $F(\mathbf{x})$, который по описанию изображения в виде вектора $\mathbf{x} = (x^1, \dots, x^S)$, состоящего из значений мультиблочных локальных шаблонов, относит это изображение либо к классу «глаз» e , либо к классу «не глаз» \bar{e} :

$$F(\mathbf{x}) = \begin{cases} +1, & \text{если } \mathbf{x} \in e \\ -1, & \text{если } \mathbf{x} \in \bar{e}. \end{cases} \tag{3}$$

Поиск итогового сильного классификатора $F(\mathbf{x})$ осуществляется с помощью алгоритма бустинга в виде суперпозиции T слабых классификаторов $f_i(\mathbf{x})$ по формуле

$$F(\mathbf{x}) = \text{sign}\left(\sum_{i=1}^T f_i(\mathbf{x})\right) \tag{4}$$



на основе обучающей выборки $(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_N, y_N)$ из N изображений, представленных в виде вектора признаков \mathbf{x}_i и метки класса y_i . Значение $y_i = +1$ соответствует изображению «глаза», а $y_i = -1$ – изображению «не глаза».

На каждой стадии t необходимо найти все возможные слабые классификаторы $f^k(\mathbf{x})$, которые ищутся в виде деревьев решений [19]:

$$f^k(\mathbf{x}) = f^k(x^1, \dots, x^k, \dots, x^S) = \begin{cases} a_0, & \text{если } x^k = 0 \\ \dots \\ a_j, & \text{если } x^k = j \\ \dots \\ a_{255}, & \text{если } x^k = 255. \end{cases} \quad (5)$$

Коэффициенты a_j вычисляются по всей обучающей выборке согласно соотношению:

$$a_j = \frac{\sum_i^N w_i y_i \delta(x_i^k = j)}{\sum_i^N w_i \delta(x_i^k = j)}, \quad (6)$$

где w_i – веса элементов выборки на шаге t . Начальные веса элементов выборки, принадлежащих классу e , выбирались больше, так как количество экземпляров «глаз» в использованной обучающей выборке в несколько раз превышало число экземпляров «не глаз».

Из всех возможных слабых классификаторов $f^k(\mathbf{x})$ на шаге t выбирается классификатор $f_t(\mathbf{x})$ с минимумом среднеквадратической ошибки:

$$f_t(\mathbf{x}) = \min_{f^k \in F} \sum_{i=1}^N w_i (y_i - f^k(\mathbf{x}_i))^2. \quad (7)$$

После этого элементам выборки присваиваются новые веса согласно соотношению:

$$w_i' = w_i e^{-y_i f_t(\mathbf{x}_i)}, \quad (8)$$

что позволяет увеличить вес сложных примеров, неверно классифицированных на данном шаге. Стартовые веса и веса на каждом шаге нормируются согласно условию:

$$\sum_{i=1}^N w_i = 1. \quad (9)$$

Так как каждый слабый классификатор, отобранный на каждой из t стадий, требует для вычисления только одного МБ-ЛБШ, значение функции $F(\mathbf{x})$ может быть найдено с помощью T отобранных МБ-ЛБШ. Таким образом, количество необходимых признаков для оценки схожести некоторого изображения с изображением глаза сокращается с величины S (порядка нескольких тысяч) до T (около 10–20).

Итерационный алгоритм локализации центров глаз

Обучим три разных классификатора $F'(\mathbf{x})$, $F''(\mathbf{x})$, $F'''(\mathbf{x})$ согласно описанной выше схеме для трех масштабов изображения глаз (рис. 4). Каждый классификатор является суперпозицией слабых классификаторов и представляется в виде выражения (4). Опуская в нем функцию $sign$, получим функцию схожести некоторого изображения с изображением глаза. Чем выше значение, тем более похоже изображение, представленное вектором признаков \mathbf{X} , на изображение глаза.

Предлагаемый новый алгоритм локализации определяет положение глаза согласно итерационному подходу, позволяющему на каждой стадии улучшать оценку местоположения центра глаза и подстраиваться к качеству и сложности изображения лица (рис. 5).

В начале работы алгоритма изображение лица представляется в грубом масштабе. Внутри области поиска центра глаза для каждого пикселя вычисляется значение функции $F'(\mathbf{x})$. Отбирается m_1 пикселей с наибольшими значениями этой функции. Выбранные точки являются грубой оценкой положения глаза.

Если разрешение изображения лица является очень низким (линейный размер лица h меньше заданного порога H'_{\min}) или функция $F'(\mathbf{x})$ имеет низкое значение для последнего из отобранных пикселей $F'_{m_1}(\mathbf{x}) < \Pi'$, то это говорит о крайне невысоком качестве изображения или о наличии на лице очков со значительными бликами или перекрытием глаз. В этом случае работа алгоритма прекращается, и в качестве положения глаза выбирается пиксель с максимальным значением $F'(\mathbf{x})$. Если же качество и размер изображения лица являются удовлетворительными, происходит переход ко второй ступени алгоритма. Здесь изображение масштабируется, для отобранных пикселей и соседних с ними в новом масштабе считается значение функции $F''(\mathbf{x})$. Отбирается m_2 пикселей с наибольшими значениями этой функции. Выбранные пиксели являются более точной оценкой положения глаза.

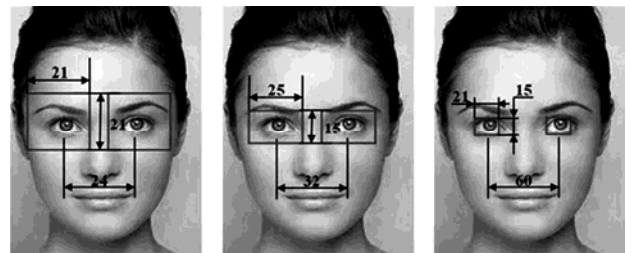


Рис. 4. Три масштаба изображения глаз для построения классификаторов $F'(\mathbf{x})$, $F''(\mathbf{x})$, $F'''(\mathbf{x})$

Если изображение лица довольно качественное и имеет относительно высокое разрешение, в качестве итогового решения о положении центра глаза выбирается пиксель с наивысшим значением функции $F'''(\mathbf{x})$, в противном случае – функции $F''(\mathbf{x})$.

При работе с изображениями лиц очень высокого разрешения число стадий и соответствующих классифи-

каторов $F(x)$ можно увеличить, что позволит определить положение центра глаза с еще большей точностью. Для этого надо предварительно обучить классификаторы для новых масштабов изображений глаз.

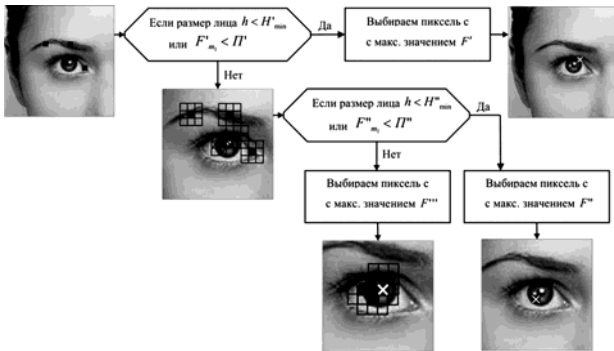


Рис. 5. Блок-схема работы

итерационного алгоритма локализации центров глаз

Результаты моделирования

Предлагаемый в работе алгоритм на основе мультиблочных локальных бинарных шаблонов, а также байесовский и градиентный алгоритмы локализации центров глаз [20, 21] программно реализованы и протестированы на базах изображений BioID [22] и FERET [23]. Для обучения байесовского локализатора и классификаторов, используемых в алгоритме на основе МБ-ЛБШ, из 3363 фронтальных изображений базы FERET с подписанными координатами левого и правого глаз выбиралась 1000 изображений. Для градиентного алгоритма локализации обучение не требуется. Для оставшихся 2363 изображений базы FERET производилось детектирование лиц стандартным алгоритмом Виолы-Джонса. Он сработал корректно на 2350 изображениях, которые и использовались для тестирования алгоритмов.

База BioID состоит из 1521 изображения с подписанными координатами глаз. Для тестирования на ней использовались алгоритмы, ранее обученные на первой тысяче изображений базы FERET, а также градиентный алгоритм, не требующий обучения. На 1469 изображениях из базы детектор лиц сработал корректно. На этих изображениях и проводилось тестирование алгоритмов локализации глаз.

В качестве критерия оценки точности локализации положений центров глаз используется нормализованная ошибка [17]:

$$err = \frac{\max(\|l - l_g\|, \|r - r_g\|)}{\|l_g - r_g\|}, \quad (10)$$

где l_g и r_g – координаты центров левого и правого глаза соответственно, определенные человеком вручную, а l и r – координаты, найденные с помощью алгоритма.

Как уже отмечалось, неточная локализация глаз может привести к существенному росту ошибок распознавания на последующем этапе обработки изображения. Тесты на стандартной базе FERET показали, что ошибки локализации $err > 0,15$ приводят к снижению уровня верного распознавания с 98% до 50% и ниже, поэтому такие ошибки будем считать грубыми. От алгоритма локализа-

ции требуется избегать таких ошибок.

На рис. 6 и рис. 7 представлены графики сравнения алгоритмов при тестировании на базах FERET и BioID соответственно. По оси абсцисс отсчитывается нормализованная ошибка, по оси ординат – доля изображений в базе, на которых алгоритм локализации глаз дал ошибку, меньшую соответствующего значения на оси абсцисс. Например, доля изображений, равная 0,96, при $err = 0,15$ показывает, что на 96% изображений из базы ошибка локализации глаз не превышает показателя 0,15.

Анализ зависимостей показывает, что разработанный итерационный алгоритм на основе МБ-ЛБШ практически не дает грубых ошибок локализации ($err > 0,15$, что соответствует смещению в среднем на половину ширины глаза от истинного положения). Только на 1% изображений из базы FERET и на 4% из базы BioID ошибка при локализации глаз превышает 0,15. Байесовский локализатор с аналогичной ошибкой работает на 2% изображений из базы FERET и 12% изображений из базы BioID. Для градиентного локализатора результаты еще хуже: 10% и 17% соответственно.

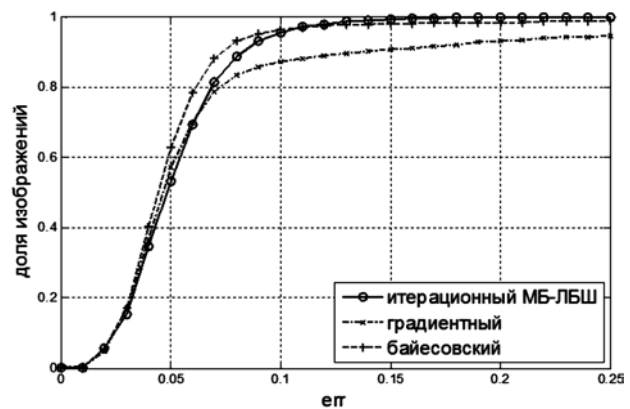


Рис. 6. Результаты тестирования алгоритмов локализации центров глаз на базе FERET

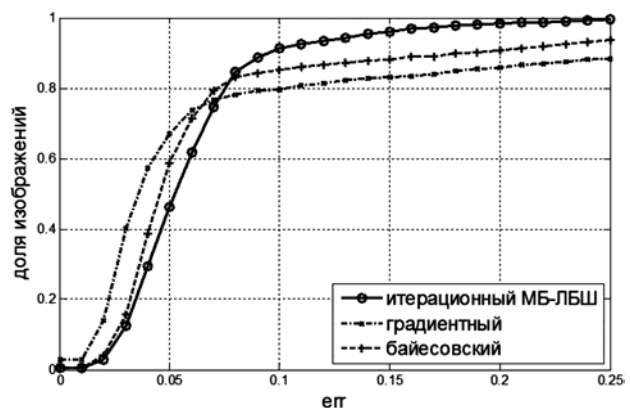
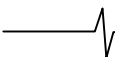


Рис. 7. Результаты тестирования алгоритмов локализации центров глаз на базе BioID

Снижение числа грубых ошибок разработанного алгоритма локализации на основе МБ-ЛБШ происходит за счет первичной оценки местоположения глаза классификатором $F'(x)$ в самом крупном масштабе. За счет адаптивной подстройки к изображениям низкого качества преимущество нового метода особенно заметно на изображениях из более сложной базы BioID. Использование



классификаторов $F''(x)$ и $F'''(x)$ позволяет новому алгоритму иметь сопоставимые результаты с другими методами локализации на простых изображениях, где нормализованная ошибка всех алгоритмов не превышает 0,1.

Результаты тестирования производительности алгоритмов представлены в табл. 1. В качестве параметра сравнения вычислялось среднее время поиска глаз на лице из базы FERET (среднее разрешение изображения лица 170x170 пикселей).

Таблица 1. Сравнение алгоритмов локализации глаз по производительности

Алгоритм	Время поиска глаз на изображении лица
Градиентный	587 мс
Байесовский	367 мс
Итерационный МБ-ЛБШ	44 мс

Видно, что новый алгоритм практически на порядок превосходит по скорости работы другие алгоритмы локализации, позволяя производить поиск положения глаз в видеопотоке в режиме реального времени.

Заключение

Предложен итерационный алгоритм локализации центров глаз на основе мультиблочных локальных бинарных шаблонов, адаптирующийся к качеству и сложности изображения лица. Разработанный алгоритм практически не дает грубых ошибок локализации ($err > 0,15$). Только на 1% изображений из базы FERET и на 4% из базы BioID ошибка при локализации глаз превышает 0,15. Рассматриваемый итерационный алгоритм практически на порядок превосходит по скорости работы другие используемые алгоритмы локализации глаз, позволяя производить поиск положения глаз в видеопотоке в режиме реального времени.

Работа выполнена при поддержке проекта № 1060 в рамках базовой части государственного задания на НИР ЯРГУ им. П.Г. Демидова и грантов РФФИ (№ 15-07-08674 и № 15-08-99639).

Литература

1. Дворкович В.П., Дворкович А.В. Цифровые видеоинформационные системы (теория и практика). – М.: Техносфера, 2012. – 1009 с.
2. Форсайт Д.А., Понс Д. Компьютерное зрение. Современный подход // М.: «Вильямс», 2004.
3. Szeliski R. Computer Vision: Algorithms and Applications – Springer, 2010.
4. Алпатов Б.А., Муравьев В.С., Стротов В.В., Фельдман А.Б. Исследование эффективности применения алгоритмов анализа изображений в задаче навигации беспилотных летательных аппаратов // Цифровая обработка сигналов. 2012. № 3. С. 29–34.
5. Алпатов Б.А., Бабаян П.В., Смирнов С.А., Масленников Е.А. Алгоритм предварительного оценивания пространственной ориентации объекта с помощью дескриптора внешнего контура // Цифровая обработка сигналов. 2014. № 3. С. 43–46.
6. Никитин А.Е., Хрящев В.В., Приоров А.Л., Матвеев Д.В. Разработка и анализ алгоритма распознавания лиц на основе локальных квантованных шаблонов // Нели-

нейный мир. 2014. № 8. С. 35–42.

7. Kriegman D., Yang M.H., Ahuja N. Detecting faces in images: A survey // IEEE Trans. on Pattern Analysis and Machine Intelligence. 2002. V. 24. № 1. P. 34–58.

8. Hjelmas E. Face detection: A Survey // Computer vision and image understanding. 2001. V. 83. № 3. P. 236–274.

9. Zhao W., Chellappa R., Phillips P., Rosenfeld A. Face recognition: A literature survey // ACM Computing Surveys (CSUR). 2003. V. 35, № 4. P. 399–458.

10. Marques J., Orlans N.M., Piszcz A.T. Effects of eye position on eigenface- based face recognition scoring // Technical Paper of the MITRE Corporation. October 2000. 7 p.

11. Никитин А.Е., Степанова О.А., Студенова А.А., Хрящев В.В. Локализация положений центров глаз на изображении лица // Докл. 17-й межд. конф. «Цифровая обработка сигналов и ее применение» (DSPA-2015). Москва, 2015. Т. 2. С. 719–723.

12. Riopka T., Boulton T. The eyes have it // Proc. of the ACM SIGMM Multimedia Biometrics Methods and Applications Workshop. 2003. P. 9–16.

13. Zhu Z., Fujimura K., Ji Q. Real-time eye detection and tracking under various light conditions // Proc. of the Symposium on Eye Tracking Research and Applications. 2002. V. 25. P. 139–144.

14. Zhu Z., Ji Q., Robust real-time eye detection and tracking under variable lighting conditions and various face orientations // Computer Vision and Image Understanding. 2005. 98 (1). P. 124–154.

15. Wang P., Green M., Ji Q., Wayman J. Automatic eye detection and its validation // Proc. of the IEEE Conference on Computer Vision and Pattern Recognition. 2005. V. 3. P. 164–172.

16. Li G. An Efficient Face Normalization Algorithm Based on Eyes Detection // Proc. of the IEEE International Conference on Intelligent Robots and Systems. 2006. P. 3843–3848.

17. Song F., Tan X., Chen S., Zhou Z.H. A literature survey on robust and efficient eye localization in real-life scenarios // Pattern Recognition. 2013. V. 46(12). P. 3157–3173.

18. Ефимов И.Н. Локальные бинарные шаблоны медианного пикселя – эффективные информативные признаки технологии распознавания образов // Цифровая обработка сигналов. 2015. № 1. С. 61–65.

19. Zhang L., Chu R., Xiang S., Liao S., Li S.Z. Face Detection Based on Multi-Block LBP Representation // Advances in Biometrics, Lecture Notes in Computer Science. 2007. P. 11–18.

20. Timm F., Barth E. Accurate Eye Centre Localisation by Means of Gradients // Proc. of the International Conference on Computer Theory and Applications (VISAPP), 2011. V. 1. P. 125–130.

21. Everingham M.R., Zisserman A. Regression and classification approaches to eye localization in face images // IEEE International Conference on Automatic Face & Gesture Recognition. 2006. P. 441–446.

22. BioID face database // URL: <https://www.bioid.com/About/BioID-Face-Database>.

23. Phillips P.J., Moon H., Rauss P.J., Rizvi S. The FERET evaluation methodology for face recognition algorithms // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2000. V. 22(10). P. 1090–1104.

УДК 004.932

МЕТОД СЖАТИЯ ЦИФРОВЫХ ИЗОБРАЖЕНИЙ БЕЗ СПЕКТРАЛЬНЫХ ПРЕОБРАЗОВАНИЙ

Петров Е.П., д.т.н., профессор, заведующий кафедрой радиоэлектронных средств Вятского государственного университета (ВятГУ), e-mail: EPPetrov@mail.ru;

Харина Н.Л., к.т.н., доцент кафедры радиоэлектронных средств ВятГУ, e-mail: NataL_res@mail.ru;

Сухих П.Н., аспирант кафедры радиоэлектронных средств ВятГУ, e-mail: raznodel@yandex.ru.

METHOD OF DIGITAL IMAGES COMPRESSION WITHOUT SPECTRAL TRANSFORMATIONS

Petrov E.P., Kharina N.L., Sukhikh P.N.

Nowadays some applications demand new methods of digital images (DI) compression which do not concede by the efficiency to the well-known methods (JPEG, JPEG2000), but demand smaller computing resources. In this work a method of compression of DI meeting the specified requirements is offered. The method provides division of a digital image into binary images (BI). Each BI is considered as a two-dimensional Markov process with two states. On the basis of the theory of conditional Markov processes a prediction of BI elements is carried out. For BIs containing low order bits preliminary separation and the removal of areas containing the background similar with the white gaussian noise (WGN) is carried out. These areas are filled with WGN selections in case of restoration. The method does not concede to the traditional methods of compression (JPEG) as for the quality/compression ratio with minimum of computing resources. It allows to work with multidigital images (8 and more) without increase in time of compression due to parallel processing of the BIs.

Keywords: digital images, image compression, binary image, Markov process, matrix of transitions probabilities, statistical characteristics, casual background, white Gaussian noise, coding of images.

Ключевые слова: цифровые изображения, сжатие изображений, марковский процесс, матрица вероятностей переходов, статистические характеристики, случайный фон, белый гауссовский шум, кодирование изображений.

Введение

В настоящее время использование цифровых изображений (ЦИ) в различных сферах деятельности, таких как телевидение, медицина, геодезия, картография, системы мониторинга и т.п. привело к необходимости использования различных методов сжатия ЦИ. Существующие методы сжатия, такие как H264, H265, JPEG, JPEG2000 успешно справляются с этой задачей [1-4]. Однако существует ряд приложений, где использование этих методов не всегда возможно из-за ограничений на вычислительные и энергетические ресурсы, например, системы дистанционного зондирования Земли (ДЗЗ), формирующие на борту космического аппарата (КА) огромные объемы информационных потоков, значительная часть которых связана с передачей цифровых изображений (ЦИ). Необходимость применения компрессии видеoinформации на борту КА обусловлена использованием датчиков, обеспечивающих съемку земной поверхности в различных спектральных диапазонах с максимально возможным пространственным разрешением. В настоящее время для кодирования ЦИ используются различные алгоритмы, базирующиеся на следующих основных методах: BTC (усеченное блочное

В настоящее время в некоторых приложениях необходимы новые методы сжатия цифровых изображений (ЦИ), не уступающие по своей эффективности известным методам (JPEG, JPEG2000), но требующие меньших вычислительных ресурсов. Предлагается метод сжатия ЦИ, удовлетворяющий указанным требованиям. Метод основан на разделении ЦИ на разрядные двоичные изображения (РДИ) и представлении РДИ двумерной цепью Маркова с двумя состояниями. На основе теории условных марковских процессов выполняется предсказание состояний пикселей РДИ. Для младших РДИ выполняется процедура удаления фона, близкого по своим статистическим характеристикам к белому гауссовскому шуму (БГШ). Эти области заполняются выборками БГШ при восстановлении. Предложенный метод не требует спектральных преобразований и позволяет осуществлять сжатие ЦИ любой разрядности (8 и более на цветовую плоскость) без увеличения вычислительных ресурсов за счет возможности параллельной обработки.

кодирование), ДИКМ (дифференциальная импульсно-кодовая модуляция), DCT (дискретное косинусное преобразование), DWT (дискретное вейвлет-преобразование) [1, 2]. Большинство современных КА обладают энергетическими возможностями для обеспечения непрерывного режима передачи информации, которые требуют больших вычислительных ресурсов на борту КА. Намечившаяся в последнее время тенденция на «миниатюризацию» космических средств наблюдения (малые спутники) ведет к ужесточению требований к эффективности использования энергетических ресурсов, и не всегда позволяет реализовать вышеперечисленные методы сжатия ЦИ, особенно ЦИ с большим разрешением.

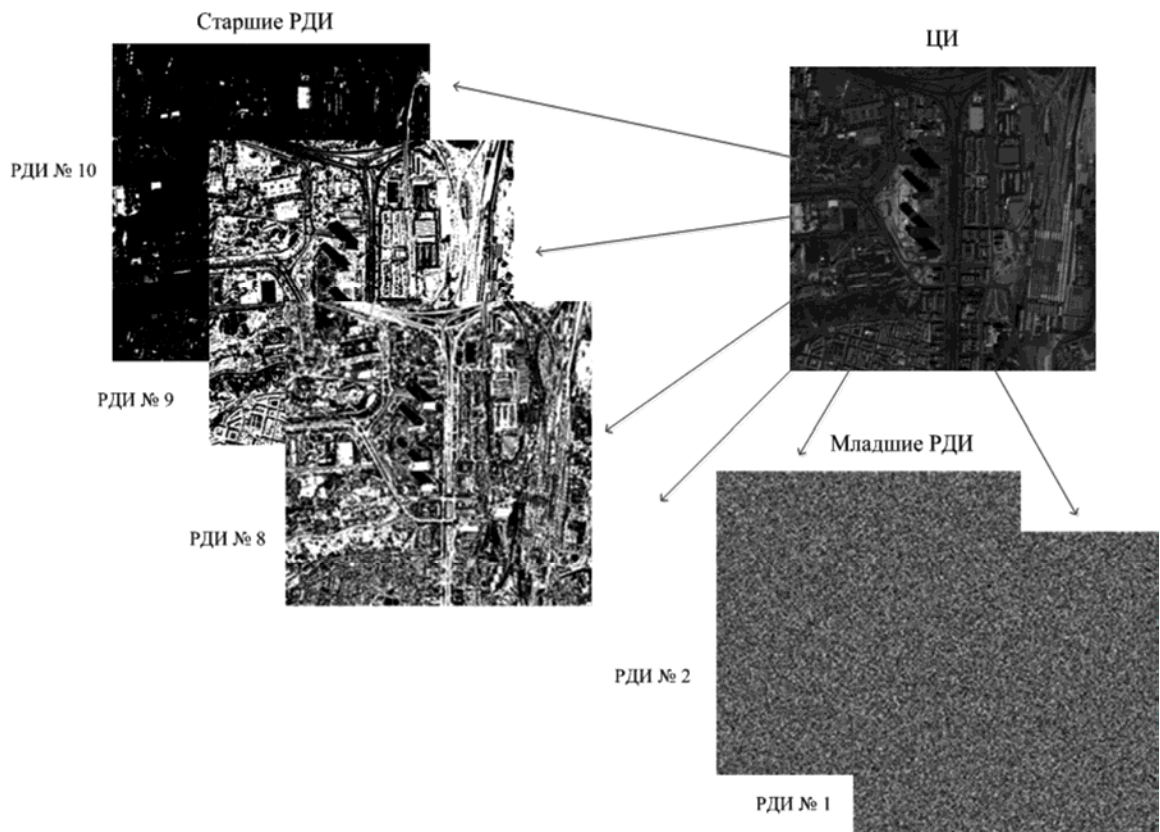


Рис. 1. Разложение ЦИ на РДИ

Отсюда следует, что для малых КА необходимы новые методы сжатия ЦИ, не уступающие по своей эффективности известным методам, но требующие меньших вычислительных ресурсов. В данной работе предлагается один из возможных методов сжатия ЦИ, удовлетворяющий указанным требованиям.

Математическая модель цифровых изображений

Будем полагать, что g -разрядное ЦИ является двумерным марковским процессом с несколькими состояниями $N = 2^g$ с вектором вероятностей начальных состояний $P = \|p_1, p_2, \dots, p_N\|^T$ и матрицами вероятностей переходов (МВП) ${}^1\Pi$ и ${}^2\Pi$ из состояния M_i в соседнее состояние M_j ($i, j \in N$) по горизонтали и вертикали соответственно [3]:

$$\begin{aligned}
 {}^1\Pi &= \begin{pmatrix} {}^1\pi_{11} & {}^1\pi_{12} & \dots & {}^1\pi_{1N} \\ {}^1\pi_{21} & {}^1\pi_{22} & \dots & {}^1\pi_{2N} \\ \dots & \dots & \dots & \dots \\ {}^1\pi_{N1} & {}^1\pi_{N2} & \dots & {}^1\pi_{NN} \end{pmatrix}, \\
 {}^2\Pi &= \begin{pmatrix} {}^2\pi_{11} & {}^2\pi_{12} & \dots & {}^2\pi_{1N} \\ {}^2\pi_{21} & {}^2\pi_{22} & \dots & {}^2\pi_{2N} \\ \dots & \dots & \dots & \dots \\ {}^2\pi_{N1} & {}^2\pi_{N2} & \dots & {}^2\pi_{NN} \end{pmatrix}.
 \end{aligned}
 \tag{1}$$

Разработка метода сжатия ЦИ в таком представлении требует больших вычислительных ресурсов, особенно для ЦИ с большим разрешением ($g > 8$). Поэтому

предлагается разделить g -разрядное ЦИ на g разрядных двоичных изображений (РДИ). На рис. 1 представлено панхроматическое 10-разрядное изображение, полученное с борта КА, после разделения на РДИ.

Каждое РДИ представляет собой двумерную цепь Маркова с двумя ($N = 2$) равновероятными состояниями ($p_1 = p_2$) и МВП ${}^1\Pi$ и ${}^2\Pi$ из состояния M_i в соседнее состояние M_j ($i, j \in N$) по горизонтали и вертикали соответственно [3]:

$${}^1\Pi = \begin{pmatrix} {}^1\pi_{11} & {}^1\pi_{12} \\ {}^1\pi_{21} & {}^1\pi_{22} \end{pmatrix}, \quad {}^2\Pi = \begin{pmatrix} {}^2\pi_{11} & {}^2\pi_{12} \\ {}^2\pi_{21} & {}^2\pi_{22} \end{pmatrix},
 \tag{2}$$

Элементы МВП (2) удовлетворяют условию нормировки

$$\sum_{j=1}^N {}^q\pi_{ij} = 1, \quad i \in N, \quad q = \overline{1, 2}
 \tag{3}$$

и стационарности

$$p_i = \sum_{j=1}^N p_j \pi_{ij}, \quad i \in N
 \tag{4}$$

РДИ представляет марковское случайное поле размером $m \times n$ на несимметричной полуплоскости, полученной классической разверткой с левого верхнего угла слева направо вниз (рис. 2).

Состояние элемента $M_{i,j}$ ($i \in m, j \in n$) зависит только от известных элементов некоторого подмножества $\Lambda_{i,j}$, называемого окрестностью элемента $M_{i,j}$. Лучшим образом удовлетворяющее условию каузальности является конфигурация окрестности (рис. 3) [5].

$$\Lambda_{i,j} = \{M_{i,j-1}, M_{i-1,j}, M_{i-1,j-1}\}. \quad (5)$$

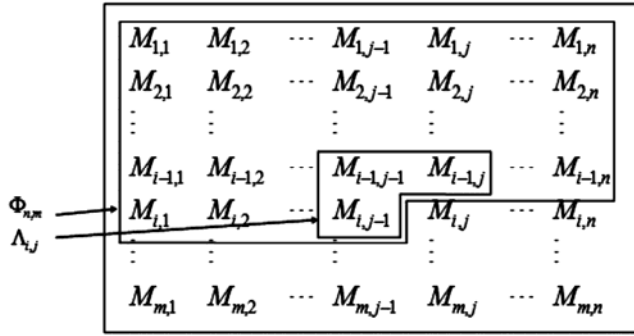


Рис. 2. Модель РДИ

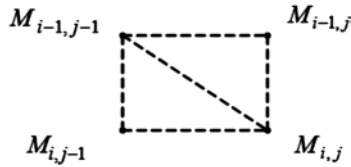


Рис. 3. Окрестность предсказываемого элемента РДИ

Вероятности перехода от состояний элементов окрестности $\Lambda_{i,j}$ к состоянию $M_{i,j}$ образуют МВП вида [5]:

$$\Pi = \begin{bmatrix} \pi_{111} & \pi_{121} & \pi_{211} & \pi_{221} \\ \pi_{112} & \pi_{122} & \pi_{212} & \pi_{222} \end{bmatrix}^T. \quad (6)$$

Элементы матрицы Π (6) связаны с элементами МВП (2) следующими соотношениями:

$$\begin{aligned} \pi_{111} &= \frac{{}^1\pi_{11} \cdot {}^2\pi_{11}}{{}^3\pi_{11}}, & \pi_{112} &= \frac{{}^1\pi_{12} \cdot {}^2\pi_{12}}{{}^3\pi_{11}}, \\ \pi_{121} &= \frac{{}^1\pi_{11} \cdot {}^2\pi_{21}}{{}^3\pi_{12}}, & \pi_{122} &= \frac{{}^1\pi_{12} \cdot {}^2\pi_{22}}{{}^3\pi_{12}}, \\ \pi_{211} &= \frac{{}^1\pi_{21} \cdot {}^2\pi_{11}}{{}^3\pi_{21}}, & \pi_{212} &= \frac{{}^1\pi_{21} \cdot {}^2\pi_{12}}{{}^3\pi_{21}}, \\ \pi_{221} &= \frac{{}^1\pi_{21} \cdot {}^2\pi_{21}}{{}^3\pi_{22}}, & \pi_{222} &= \frac{{}^1\pi_{22} \cdot {}^2\pi_{22}}{{}^3\pi_{22}}. \end{aligned}$$

где ${}^3\pi_{ii}$ – элементы дополнительной МВП ${}^3\Pi = {}^1\Pi \times {}^2\Pi'$.

Метод сжатия цифровых изображений

Основой алгоритма является процедура предсказания состояния каждого пиксела РДИ на основе известной окрестности $\Lambda_{i,j}$ и МВП (6), используя критерий максимальной вероятности. В результате предсказания формируется РДИ по следующему правилу – если состояние пиксела предсказано верно, то этот пиксел в РДИ принимает состояние 0, иначе 1. В качестве примера на рис. 4 а, б, в, г приведены результаты предсказания двух старших РДИ ЦИ, белым показаны области, присутствующие в битовом потоке (рис. 3 б). При восстановлении РДИ нулевые пикселы заменяются состояниями по результатам предсказания, а единичные – неправильно предсказанные – инверсией результата предсказания.

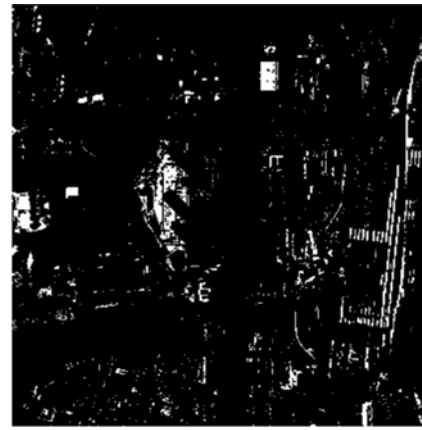


Рис. 4а. РДИ № 10



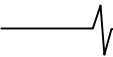
Рис. 4б. РДИ № 10 после процедуры предсказания



Рис. 4в. РДИ № 9



Рис. 4 г. РДИ № 9 после процедуры предсказания



РДИ после предсказания представляет собой разреженное поле, что позволяет эффективно применять алгоритмы кодирования RLE и Хаффмана.

Формализованное описание алгоритма [4, 5]:

1. ЦИ разбивается на g РДИ;
2. Для каждого РДИ вычисляются МВП по горизонтали $^1\Pi$ и вертикали $^2\Pi$, соответственно;
3. Исходя из МВП $^1\Pi$ и $^2\Pi$ вычисляется МВП $^3\Pi$.
4. Прогнозируется элемент \hat{M}_{ij} на основе МВП Π и окрестности $\Lambda_{i,j}$.

5. Если $M_{ij} = \hat{M}_{ij}$, то элемент предсказан верно;

6. Если $M_{ij} \neq \hat{M}_{ij}$, элемент M_{ij} предсказан неверно;

7. Формирование РДИ после предсказания;

8. РДИ сжимается кодом RLE в одномерный поток;

9. Одномерный поток сжимается кодом Хаффмана;

10. Сжатые данные записываются в файл.

Средние и младшие РДИ (№№1-6) содержат значительные области с фоном, характеристики которого близки к белому гауссовскому шуму. Процедура предсказания в этом случае мало эффективна, поэтому из младших РДИ на основе априорно известных или вычисленных статистических характеристик осуществляется удаление случайного фона и замена его детерминированным фоном, на котором остаются детали объектов ЦПИ с резко выраженными границами. После удаления фона РДИ подвергается кодированию, т.е. выполняются пункты 8-10 алгоритма сжатия. При восстановлении области детерминированного фона заполняются выборками белого гауссовского шума.

На рис. 5а приведено РДИ № 5 и выделенные черным области детерминированного фона (рис. 5б).

С увеличением разрядности у большинства ЦИ в младших РДИ присутствует практически один фон. На рис. 6а,б представлены два младших РДИ 10-разрядного ЦПИ. Во втором РДИ обнаружены мелкие детали объектов (рис. 7), они представлены белыми областями, черным цветом обозначен детерминированный фон. В первом, самом младшем РДИ, деталей объектов не обнаружено, следовательно оно кодированию не подлежит и целиком заполняется выборками белого гауссовского шума при восстановлении. Такой эффект характерен для большинства исследованных 10-разрядных спутниковых снимков.

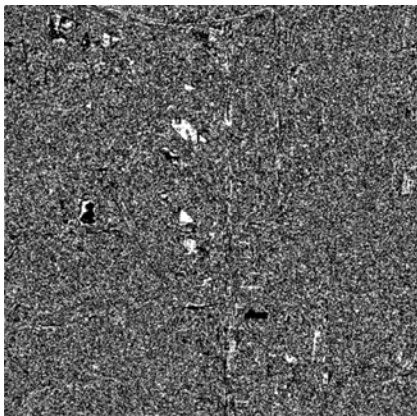


Рис. 5а. РДИ № 5



Рис. 5б. РДИ № 5 после удаления фона

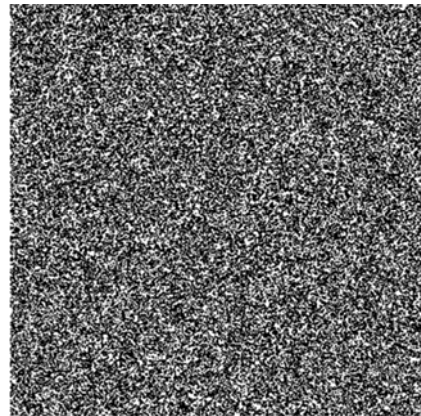


Рис. 6а. РДИ № 2

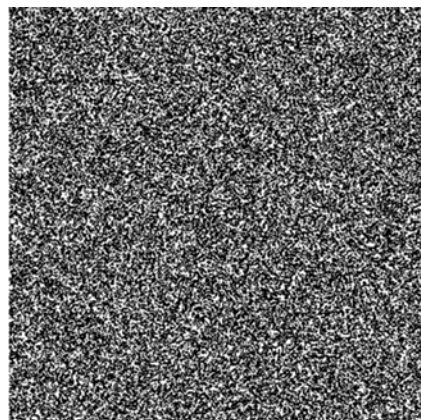


Рис. 6б. РДИ № 1



Рис.7 РДИ № 2 после удаления фона

На рис. 8а приведено 10 разрядное панхроматическое ЦПИ (фрагмент 2000×2000) и восстановленное после сжатия предложенным методом (рис. 8 б) с коэффициентом сжатия 6,3. Эффективность разработанного метода сжатия слабо зависит от разрядности ЦПИ в пределах 8-16 разрядов.



Рис. 8а. 10-разрядное панхроматическое ЦПИ



Рис. 8б. 10-разрядное ЦПИ, восстановленное после сжатия предложенным методом, коэффициент сжатия 6,3

оценкам качества показывает, что предложенный метод несколько уступает известным методам сжатия. Это происходит вследствие заполнения удаленных областей фона выборками белого гауссовского шума при восстановлении, что никак не сказывается на точности сохранения деталей объектов.

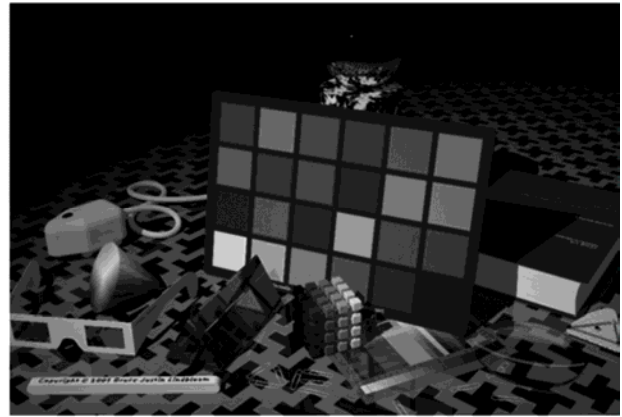


Рис. 9а. Тестовое 24-битное цветное RGB ЦИ

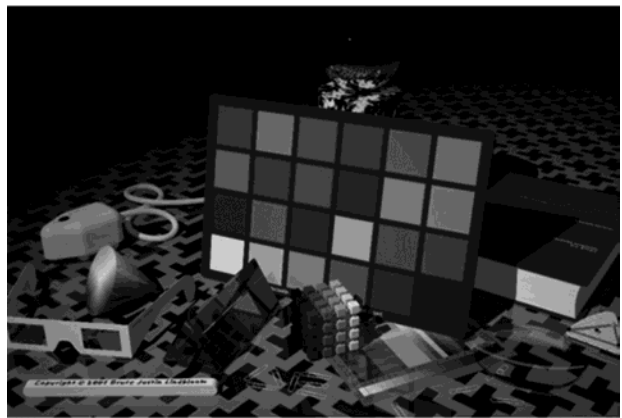


Рис. 9б. 24-битное цветное RGB ЦИ, восстановленное после сжатия предложенным методом, коэффициент сжатия 18

На рис. 9 приведено другое тестовое 24-битное цветное RGB изображение до сжатия (а) и после (б) с коэффициентом сжатия 18.

Для оценки эффективности алгоритма в сравнении с другими алгоритмами взяты параметры MSE и SSIM (рис 10 и рис. 11).

Анализ эффективности предложенного метода по

Кроме того предложенный метод обладает более высоким быстродействием по отношению к аналогам, что является важнейшим критерием в обработке изображений в системах дистанционного зондирования Земли. На рис. 12 представлены зависимости скорости сжатия без потерь тестовых изображений в мс от размера изображений (в процентах).

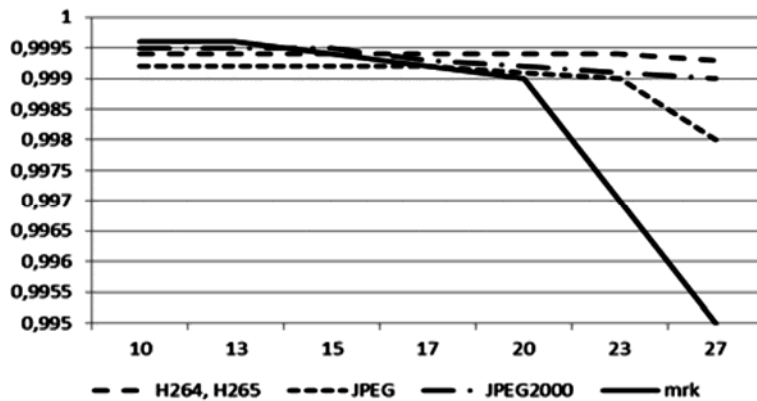


Рис. 10. График зависимости MSE от коэффициента сжатия для предлагаемого метода (формат mrk) и аналогов

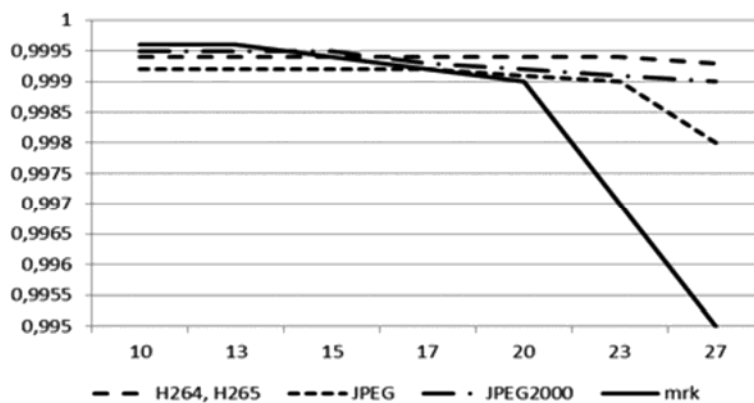


Рис. 11. График зависимости SSIM от коэффициента сжатия для предлагаемого метода (формат mrk) и аналогов

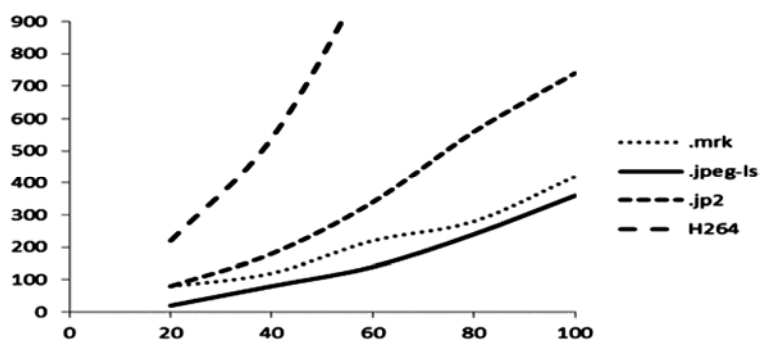


Рис. 12. Зависимости скорости сжатия без потерь тестовых изображений

Заключение

Алгоритм имеет ряд преимуществ перед другими методами сжатия:

- сжатие каждого РДИ происходит независимо от других РДИ, что дает возможность увеличить быстродействие за счет параллельной обработки РДИ;
- алгоритм позволяет обрабатывать изображения любой разрядности;
- алгоритм строится на простых вычислительных операциях, что снижает требования к элементной базе.

Литература

1. Гонсалес Р., Вудс Р. Цифровая обработка сигналов. М.: Техносфера, 2005. - 1072 с.
2. Яне Б. Цифровая обработка изображений. М.: Техносфера, 2007. – 584с.
3. Дворкович В.П., Дворкович А.П. Цифровые видеоинформационные системы (теория и практика). М.: Техносфера, 2012. – 1009с.

4. Дворкович В.П., Дворкович А.П. Метрологическое обеспечение видеоинформационных систем. М.: Техносфера, 2015. – 784с.

5. Петров Е.П., Медведева Е.В., Харина Н.Л. Математическая модель цифровых полутоновых изображений Земли из космоса // II Всероссийская НТК «Актуальные проблемы ракетно-космической техники». Самара. 2011. С. 179-185.

6. Петров Е.П., Харина Н.Л., Ржаникова Е.Д. Метод сжатия цифровых полутоновых изображений на основе цепей Маркова с несколькими состояниями// Труды РНТОРЭС им. А. С. Попова. Серия: Цифровая обработка сигналов и ее применение. – М., 2013. Выпуск: XV-1, – с. 132-135

7. Петров Е.П., Харина Н.Л., Ржаникова Е.Д. Метод сжатия цифровых полутоновых изображений на основе цепей Маркова с несколькими состояниями//Материалы III Всероссийской НТК «Актуальные проблемы ракетно-космической техники» (III Козловские чтения) – Самара, 2013. – с.163-170

УДК 004.932

ИССЛЕДОВАНИЕ АЛГОРИТМОВ КОРРЕЛЯЦИОННОГО СОВМЕЩЕНИЯ ИЗОБРАЖЕНИЙ В СИСТЕМАХ КОМБИНИРОВАННОГО ВИДЕНИЯ

Елесина С.И., доцент кафедры «Электронные вычислительные машины» ФГБОУ ВПО «Рязанский государственный радиотехнический университет», e-mail: elesinas@mail.ru;

Ломтева О.А., аспирант кафедры «Электронные вычислительные машины» ФГБОУ ВПО «Рязанский государственный радиотехнический университет», e-mail: olomteva@gmail.com.

RESEARCH OF THE CORRELATION IMAGES COMBINING ALGORITHMS IN COMBINED VISION SYSTEMS

Elesina S.I., Lomteva O.A.

The authors research global optimization method – genetic algorithm with the aim to receive its optimum parameters values while it being used in combined vision systems has been carried out. Extended angles of virtual terrain model are supposed to decrease the computational complexity of the correlation combining images algorithm. This approach considerably reduces the number of virtual images generated for combining.

Keywords: global extremum, criterion function, search area, genetic algorithm, step-by-step scanning, images pyramid, combined vision systems, parameters of aircraft positioning, real image, virtual image.

Ключевые слова: биполярный сигнал, метод Эйлера-Коши, фильтрация, нелинейный фильтр, цифровое моделирование.

Введение

Повышение безопасности и эффективности полетов на различных типах летательных аппаратов (ЛА) требует разработки новых решений в области авионики, направленных на улучшение ситуационной информированности экипажа. Вследствие этого в последнее время одной из наиболее актуальных задач авиатехники стало создание систем комбинированного видения (Combined Vision System – CVS), которые представляют собой комбинацию подсистем искусственного видения (Synthetic Vision System – SVS) и улучшенного видения (Enhanced Vision System – EVS). Важной функцией таких комбинированных систем технического зрения является совмещение реальных изображений (РИ), сформированных бортовой камерой ЛА, и виртуальных изображений (ВИ), синтезированных с учетом текущего положения ЛА на основе цифровой карты местности (ЦКМ). Результат такого совмещения гораздо более информативен для летчика, чем исходные изображения, наблюдаемые им по отдельности. В результате повышается степень ситуационной осведомленности экипажа, что ведет к снижению аварийности пилотирования, экономии топлива и ресурсов ЛА. Кроме того, совмещение реального и виртуального изображений позволяет осуществлять корректировку координат ЛА [1, 2].

И реальное, и виртуальное изображения можно охарактеризовать с помощью следующих 6 параметров позиционирования: высота h , широта φ , долгота λ , крен γ , курс ψ , тангаж θ . Для ВИ перечисленные параметры известны. Для РИ эти параметры имеют погрешности из-за неточной работы навигационного оборудования.

Проведены исследования метода глобальной оптимизации – генетического алгоритма с целью получения оптимальных значений его параметров при использовании в системах комбинированного видения. Для уменьшения вычислительной сложности данного алгоритма корреляционного совмещения изображений предлагается использовать расширенные ракурсы виртуальной модели местности. Данный подход значительно сокращает количество генерируемых для совмещения виртуальных изображений.

Таким образом, сравнивая РИ и ВИ, можно получить более точные текущие параметры позиционирования, а по ним – текущие координаты ЛА.

Общая схема совмещения изображений и коррекции координат ЛА системой комбинированного видения представлена на рис. 1 и заключается в следующем.

От сенсоров системы технического зрения (СТЗ) РИ поступает на блоки «Предварительная обработка». Работа блоков такого типа основана на том, что обрабатываемые РИ содержат избыточную информацию, устранение которой не влечет за собой снижения вероятности и точности совмещения. Сокращение исходной информации посредством снижения информативности изображений является одним из способов уменьшения объема вычислений и увеличения вероятности правильного корреляционного совмещения изображений. РИ может быть различной физической природы в зависимости от СТЗ (телевизионное, тепловизионное, радиолокационное).

С другой стороны по навигационным и пилотажным данным (параметрам позиционирования ЛА) по ЦКМ формируется ракурс виртуальной модели местности (ВММ) – виртуальное изображение (ВИ), которое также проходит предварительную обработку. Обработка изображений (реальных и виртуальных) носит совершенно разный характер для совмещения и визуализации. В первом случае это поиск общих характерных элементов, по которым реальные и виртуальные изображения мож-

но совместить; во втором – наиболее эффективный способ визуализации итогового комбинированного изображений [3].

В блоке «Совмещение» выполняется корреляционное совмещение двух РИ и ВИ, а именно вычисляется функционал сравнения, в частности, взаимно-корреляционная функция данной пары изображений, и определяется экстремум полученного функционала сравнения (рис. 2).

Наибольшую вычислительную сложность имеет подсистема поиска экстремума функционала сравнения. Причем вычислительная сложность поиска экстремума в значительной степени зависит от типа критериальной (целевой) функции (функционала сравнения). Так как реальное и виртуальное изображения являются бинарными, то наиболее предпочтительными являются парные критериальные функции. В работах [4, 5] приведен тщательный анализ и отбор лучших парных критериальных функций для решения поставленных задач. Наиболее качественной функцией, демонстрирующей наилучшие результаты, является парная функция Рао (1), и именно она используется в данной работе.

$$K = \frac{a}{a+b+c+e}, \quad (1)$$

где значение a соответствует совпадению черных пикселей на сравниваемых изображениях, e соответствует совпадению белых пикселей на сравниваемых изображениях, b и c – пиксели на изображениях, имеющие различную яркость: белую и черную, черную и белую соответственно.

Целью данной работы является исследование методов поиска глобального экстремума (ГЭ) целевой функции (ЦФ) при корреляционном совмещении РИ и ВИ, а также поиск путей повышения их быстродействия.

В соответствии с целью поставлены следующие задачи:

- 1) отбор и практическая реализация подходящих для решения рассматриваемой задачи методов поиска ГЭ ЦФ при корреляционном совмещении РИ и ВИ;
- 2) проведение серии экспериментов для исследования работы методов поиска ГЭ ЦФ и получения оптимальных настроек их параметров;
- 3) сравнение эффективности работы методов поиска ГЭ ЦФ и выбор наилучшего среди них;
- 4) дальнейшее уменьшение вычислительной сложности метода поиска ГЭ ЦФ, выбранного в качестве наилучшего.

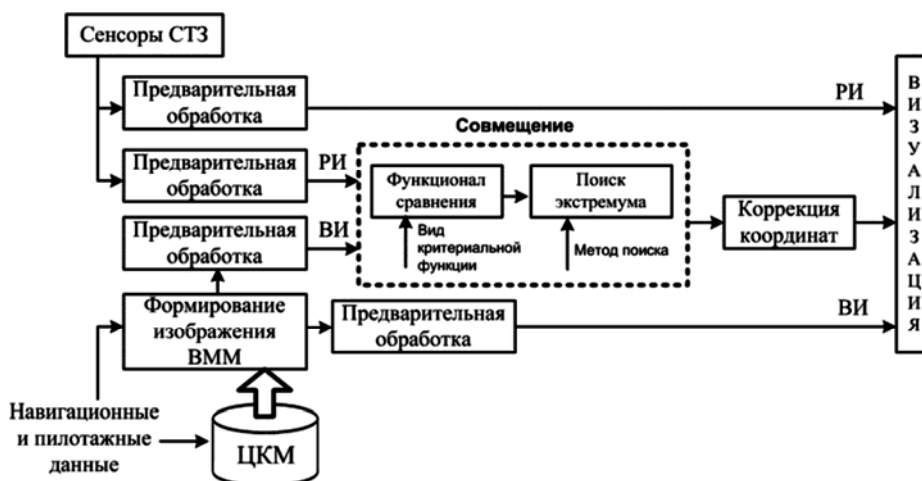


Рис. 1. Общая схема совмещения изображений и коррекции координат системой комбинированного видения ЛА

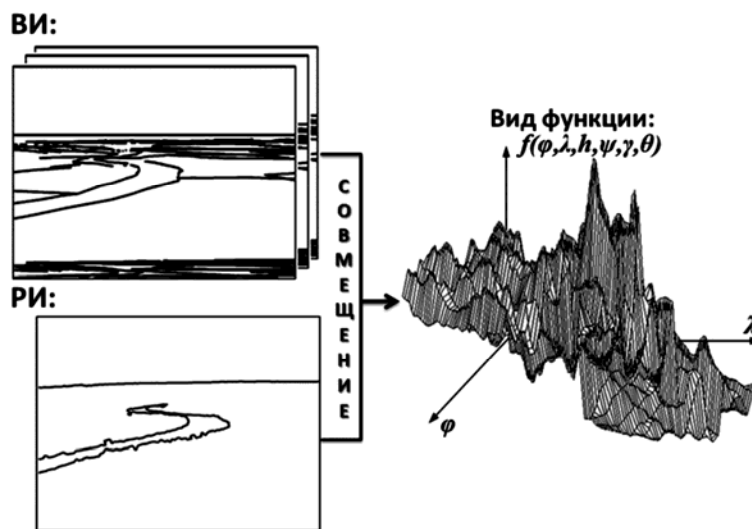


Рис.2. Вид проекции на координаты λ и φ многоэкстремальной ЦФ при совмещении изображений по шести параметрам позиционирования ЛА

Существует множество различных методов поиска ГЭ. Наиболее простым и надежным является метод полного перебора всех возможных ВИ (ракурсов ВММ) и РИ. Такой метод дает очень хорошие результаты по точности совмещения. Однако этот метод в чистом виде непригоден для использования в системах комбинированного видения, так как в общем случае приходится проводить корреляционное совмещение изображений в шестимерном пространстве D по шести степеням свободы: широта, долгота, высота, крен, курс, тангаж: $D = \varphi \cdot \lambda \cdot h \cdot \gamma \cdot \psi \cdot \theta$. Если число переборных дискретных значений по координатам равно десяти, а по углам равно пяти, то при полном переборе необходимо построить $N = 10 \cdot 10 \cdot 10 \cdot 5 \cdot 5 = 125\ 000$ вариантов ракурсов изображения ВММ и столько же раз выполнить совмещение, что невозможно в условиях реального времени.

Для сокращения трудоемкости поиска ГЭ значения каждого параметра позиционирования можно перебирать с некоторым шагом $h > 1$, а затем произвести уточнение найденного таким образом экстремума, «сканируя» его окрестность с единичным шагом по каждому измерению. Это так называемый метод *позатипного сканирования*.

Метод поэтапного сканирования можно дополнительно оптимизировать, используя «пирамиду» изображений (рис. 3), то есть работая с изображениями различных масштабов на каждом этапе сканирования (уровне пирамиды). Уменьшенный в несколько раз масштаб изображения на верхнем уровне пирамиды (первом этапе) значительно сокращает количество точек, что приводит к ускорению работы алгоритма корреляционного совмещения [6].

Перечисленные методы являются детерминированными, т.к. при поиске экстремума ни разу не используются случайные значения его параметров. Хотя данные методы и обеспечивают высокую точность нахождения ГЭ, трудоемкость их слишком велика, поскольку они пассивны, «слепо» проходят зону поиска и не учитывают никакой информации, получаемой в ходе работы.

От проблем детерминированных алгоритмов позволяют уйти стохастические алгоритмы, в работе которых присутствует элемент случайности. Примером таких алгоритмов являются эволюционные алгоритмы, в частности, их разновидность – *генетический алгоритм (ГА)*.

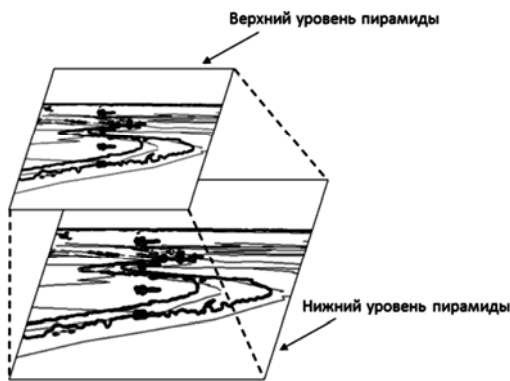


Рис. 3. Пирамида изображений

Проведенные сравнительные испытания показали, что ГА отличается высокой скоростью работы по сравнению с методом поэтапного сканирования и методом, использующим пирамиду изображений, и при этом он не уступает им по точности полученных результатов (рис. 4).



Рис. 4. Зависимость вероятности попадания в ГЭ от трудоемкости для различных методов поиска ГЭ

Поэтому для решения поставленной задачи следует применять именно ГА, а дальнейшие исследования необходимо направить на поиск оптимальных настроек данного алгоритма, а также способов уменьшения его трудоемкости. С этой целью предлагается использование расширенных ракурсов ВММ при совмещении РИ и ВИ. Они получаются путем исключения двух параметров позиционирования ЛА (тангажа и курса) и увеличения линейных размеров каждого отдельно взятого ракурса (ВИ).

Принцип построения расширенных ракурсов заключается в следующем. Поле зрения расширяется, как показано на рис. 5, по вертикали и по горизонтали с учетом интервалов неопределенности углов тангажа и курса соответственно [3].

Для расчета размера сторон расширенного ракурса ВММ (рис. 5) используются формулы (2) и (3):

$$W_x = X \cdot \frac{\operatorname{tg}\left(\frac{\alpha_0 + 2\Delta\psi}{2}\right)}{\operatorname{tg}\left(\frac{\alpha_0}{2}\right)}, \quad (2)$$

$$W_y = Y \cdot \frac{\operatorname{tg}\left(\frac{\alpha_1 + 2\Delta\theta}{2}\right)}{\operatorname{tg}\left(\frac{\alpha_1}{2}\right)}, \quad (3)$$

где X, Y – размеры реального изображения.

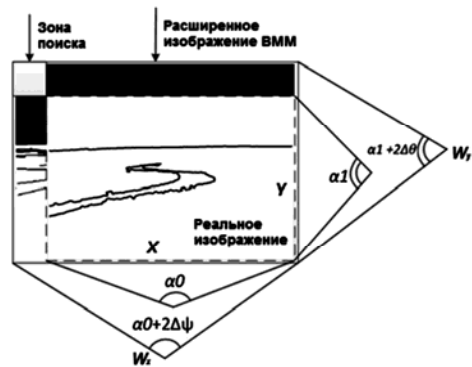


Рис. 5. Размеры расширенного изображения ВММ относительно реального изображения

Генетический алгоритм

Идеи ГА позаимствованы у природы [7, 8]. Они основаны на генетических процессах живых организмов: биологические популяции развиваются в течение нескольких поколений, подчиняясь законам естественного отбора по принципу «выживает наиболее приспособленный». Подражая этому процессу, ГА способны «развивать» решения реальных задач, закодированных соответствующим образом. ГА используют для своей работы эволюционные принципы наследственности, изменчивости и естественного отбора. Общая схема ГА в теории и в приложении к данной предметной области представлена на рис. 6.

ГА работает с некоторым случайным набором исходных решений, называемых *популяцией особей*, в хромосоме (генотип) каждой из которых закодировано возможное решение задачи (фенотип).

Для решения задачи совмещения изображений с помощью ГА её параметры необходимо представить в виде генов особи. Это значения шести параметров позиционирования ЛА: широта, долгота, высота, курс, крен и тангаж. Поскольку мы имеем дискретное пространство поиска, т.е. каждый параметр задачи может принимать лишь отдельные допустимые значения, то целесообразно выбрать вещественное кодирование хромосом. Каждая хромосома будет состоять из шести отдельных вещественных чисел (генов), которые соответствуют параметрам позиционирования конкретного ВИ (особи).

Изначально популяция формируется случайным образом (блок «Инициализация популяций решений» на рис.6). Оценивание качества закодированных решений производится с помощью функции приспособленности (ЦФ), в роли которой выступает парная критериальная

функция (КФ) Рао (блок «Оценивание особей» на рис. 6). По результатам оценивания наиболее приспособленные особи выбираются (блок «Селекция» на рис. 6) для скрещивания.

В результате скрещивания выбранных особей посредством применения генетического оператора *кроссинговера* создается потомство, генетическая информация которого формируется в результате обмена хромосомной информацией между родительскими особями (блок «Скрещивание» на рис. 6).

Созданные потомки формируют новую популяцию, причем часть потомков мутирует (используется генетический оператор *мутации*), что выражается в случайном изменении их генотипов (блок «Мутация» на рис. 6). Оператор мутации необходим для «выбивания» популяции из локального экстремума и способствует защите алгоритма от преждевременной сходимости. В данной задаче мутация изменяет значения отдельных генов (параметров позиционирования) особей на другие, случайно взятые из множества их допустимых значений.

Перечисленная последовательность действий называется *поколением*. Эволюция популяции состоит из последовательности таких поколений. Результат эволюции – определение ГЭ.

Результат работы ГА во многом зависит от настройки его параметров: количество поколений, размер популяции, тип селекции, давление селекции, разновидность оператора кроссинговера, вероятность кроссинговера P_c , разновидность оператора мутации, вероятность мутации P_m .

Для уменьшения трудоемкости ГА используется кластеризация области поиска [5, 9]. Допустимый размер кластера определяется расстоянием между локальными экстремумами.



Рис. 6. Общая схема ГА

Экспериментальные исследования

При генерации ракурсов ВММ в ходе исследования работы ГА приняты следующие пределы и шаги изменения поправок к навигационным данным:

- по широте φ и долготе λ : в интервале $[-32; 32]$ с шагом 16
(-32, -16, 0, 16, 32) – 5 возможных значений;
- по высоте h : в интервале $[-32; 32]$ с шагом 8
(-32, -24, -16, -8, 0, 8, 16, 24, 32) – 9 возможных значений;
- по крену γ : в интервале $[-0.2; 0.2]$ с шагом 0.2
(-0.2, 0, 0.2) – 3 возможных значения;
- по курсу ψ : в интервале $[-1; 1]$ с шагом 0.5
(-1, -0.5, 0, 0.5, 1) – 5 возможных значений;
- по тангажу θ : в интервале $[-0.2; 0.2]$ с шагом 0.1
(-0.2, -0.1, 0, 0.1, 0.2) – 5 возможных значений.

При этом общее число вариантов ракурсов равно $5 \cdot 5 \cdot 9 \cdot 3 \cdot 5 \cdot 5 = 16875$, каждый ракурс занимает 469 Кбайт дискового пространства, а их общий размер составляет около 7,55 Гб. Таким образом, обеспечивается достаточная точность оценки навигационных параметров, и при этом объем полученной базы данных ракурсов позволяет хранить ее на жестком магнитном диске.

Определение оптимальных параметров ГА

Для нахождения оптимальных для решаемой задачи настроек ГА были проведены эксперименты, разбитые на серии в зависимости от параметров ГА. Число экспериментов в каждой серии – 300. Размеры РИ и ВИ одинаковы и составляют 800×600 пикселей.

На рис. 7 представлена зависимость вероятности попадания в ГЭ от количества поколений и размера популяции.

Исходя из данного рисунка, можно сделать вывод, что в дальнейшем целесообразно выбирать размер популяции не менее 500 особей и количество поколений не менее 20, чтобы вероятность попадания в ГЭ превышала 0,95.

Исследования показали, что чем больше вероятность скрещивания и мутации, тем лучшие результаты показывает алгоритм. Для достижения попадания в глобальный экстремум с вероятностью больше 0,95 необходимо выбирать P_c и P_m , равными 0,1-0,3 и 0,03-0,05 соответственно.

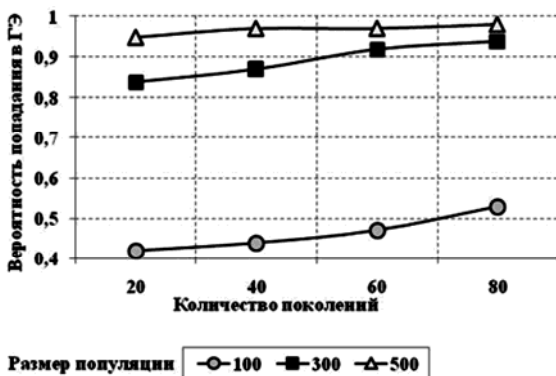


Рис. 7. Зависимость вероятности попадания в ГЭ от количества поколений и размера популяции

Результаты экспериментов для трех видов селекции приведены на рис.8 (при размере популяции 500, количестве поколений 20, $P_c = 0,25$ и $P_m = 0,05$).

Из диаграммы видно, что худшие показатели имеет рулеточная селекция (вероятность не более 0,87), а лучшие – селекция усечением (вероятность до 0,95).

Результат совмещения реального и виртуального изображений методом ГА представлен на рис. 9.

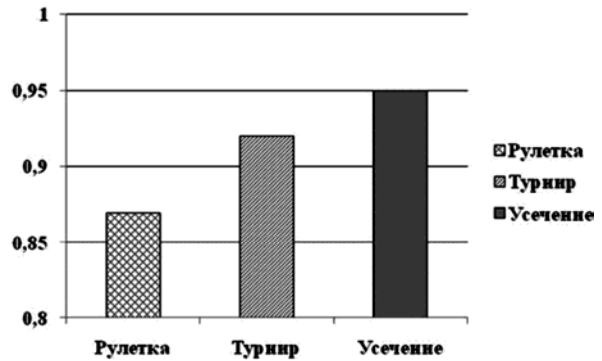


Рис. 8. Зависимость вероятности попадания в ГЭ от различных типов селекции



Рис. 9. Результат совмещения РИ и ВИ с использованием ГА

Исследование эффективности применения расширенных ракурсов ВММ

Размеры сторон расширенных ракурсов ВММ вычислялись по формулам (2) и (3). В качестве погрешностей $\Delta\psi$ и $\Delta\theta$ использовались предельно возможные ошибки по данным координатным осям, соответствующие 1° и $0,2^\circ$ соответственно. Углы зрения телекамеры α_1 и α_0 равны $20^\circ \times 30^\circ$. Размер РИ равен 800×600 пикселей. При этом размер расширенного ракурса составил 856×612 пикселей.

Таким образом, размер зоны поиска РИ на расширенном ВИ (рис. 5) составил 56×12 пикселей (672 возможных вариантов). Прохождение данной зоны можно осуществлять с некоторым шагом (метод сканирования). Экспериментально установлено, что оптимальными являются следующие шаги сканирования: 8 по ширине и 2 по высоте. Тогда количество вызовов ЦФ при прохождении зоны поиска сократится в 16 раз.

На рис. 10 представлены графики зависимости вероятности попадания в ГЭ от трудоемкости (количества

построенных ВИ) ГА при использовании обычных и расширенных ракурсов ВММ. Видно, что метод использования расширенных ракурсов так же, как и первый метод, обеспечивает высокую вероятность попадания в ГЭ (около 92 %), но при этом генерируется почти в 20 раз меньше ВИ.

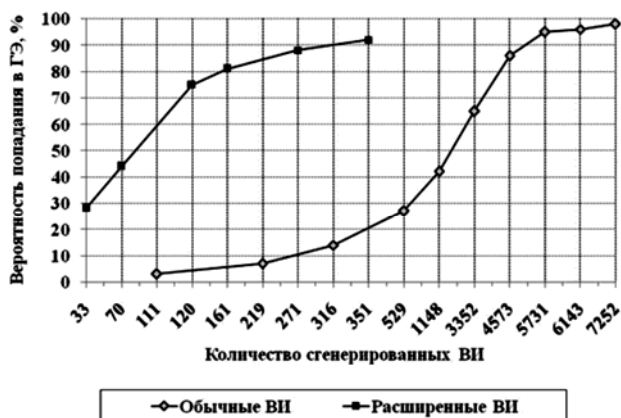


Рис. 10. Зависимость вероятности попадания в ГЭ от трудоемкости ГА при использовании обычных и расширенных ракурсов ВММ

Заключение

Исходя из полученных результатов исследования, оптимальными настройками ГА для решения задачи совмещения реального и виртуального изображений в комбинированных системах видения являются:

- 1) длительность эволюции (количество поколений): не менее 20;
- 2) размер популяции: не менее 500;
- 3) интенсивность (давление) селекции: 0,3;
- 4) разновидность оператора кроссинговера: двухточечный для вещественного кодирования;
- 5) вероятность кроссинговера P_c : 0,1-0,3;
- 6) разновидность оператора селекции: усечение;
- 7) вероятность мутации P_m : 0,03-0,05.

Использование данных параметров ГА обеспечивает достаточно высокую вероятность попадания в ГЭ: более 0,95 (в случае, если РИ и ВИ визуальны схожи).

Рассмотренный метод использования расширенных ракурсов ВММ приводит к отказу от перебора значений углов тангажа и курса при рендеринге изображения

ВММ, т.к. все изменения, вносимые этими углами при определенном наборе других навигационных параметров, включены в расширенный ракурс ВММ. В результате значительно сокращается количество используемых ракурсов ВММ и, следовательно, повышается быстродействие ГА.

Литература

1. Костяшкин Л.Н., Логинов А.А., Никифоров М.Б. Проблемные аспекты системы комбинированного видения летательных аппаратов // Известия ЮФУ. Технические науки. 2013. № 5. С. 61-65.
2. Логинов А.А., Муратов Е.Р., Никифоров М.Б., Новиков А.И. Сокращение вычислительной сложности совмещения изображений в авиационных системах технического зрения // Динамика сложных систем. 2015. Т. 9, № 1. С. 33-40.
3. S. Elesina, O. Lomteva. Increase of image combination performance in combined vision systems using genetic algorithm // Proceedings of the 3rd Mediterranean Conference on Embedded Computing. – Budva, Montenegro.– 2014. PP. 158-161.
4. M. Nikiforov, S.Elesina, A.Efimov, “Criteria Functions Selection for Combined and Enhanced Synthetic Vision Systems of the Aircraft,” Computer Science and Information Technologies: Materials of the VIIth International Scientific and Technical Conference CSIT 2013. Lviv: Publishing Lviv Polytechnic, 2013.С. 56-58.
5. Елесина С.И., Костяшкин Л.Н., Логинов А.А., Никифоров М.Б. Совмещение изображений в корреляционно-экстремальных навигационных системах. Монография // Под ред. Костяшкина Л.Н., Никифорова М.Б. – М.: Радиотехника, 2015. 208 с.
6. Визильтер Ю.В., Желтов С.Ю., Бондаренко А.В., Ососков М.В., Моржин А.В. Обработка и анализ изображений в задачах машинного зрения: Курс лекций и практических занятий. – М.: Физматкнига, 2010. 672 с.
7. Цой Ю.Р., Спицын В.Г. Генетический алгоритм: учебное пособие. Томск: Изд-во ТПУ, 2006. 146 с.
8. M. Mitchell, An Introduction to Genetic Algorithms, 5th ed. Cambridge, MA: MIT Press, 1999.С. 103-131.
9. Елесина С.И., Никифоров М.Б. Повышение эффективности генетического алгоритма // Информационные технологии: теоретический и прикладной научно-технический журнал. 2012. № 3. С. 49-54.

AUTEX Ltd.



**18-я Международная Конференция
ЦИФРОВАЯ ОБРАБОТКА СИГНАЛОВ
И ЕЕ ПРИМЕНЕНИЕ**

**The 18th International Conference DIGITAL SIGNAL PROCESSING AND ITS
APPLICATIONS**

Москва, 29 марта – 1 апреля 2016 года (ориентировочно)

Уважаемые коллеги!

ПРИГЛАШАЕМ ВАС ПРИНЯТЬ УЧАСТИЕ В РАБОТЕ КОНФЕРЕНЦИИ

ОРГАНИЗАТОРЫ:

- Российское научно-техническое общество радиотехники, электроники и связи им. А.С. Попова
- Институт радиотехники и электроники РАН
- Компания AUTEX Ltd. (ЗАО «АВТЭКС»)
- Российская секция IEEE
- IEEE Signal Processing Society
- Институт проблем управления РАН
- Институт проблем передачи информации РАН
- Московский научно-исследовательский телевизионный институт (ЗАО МНИТИ)

ОБЪЕДИНЁННЫЙ ОРГАНИЗАЦИОННЫЙ КОМИТЕТ КОНФЕРЕНЦИИ:

Председатель Международного комитета: академик **Гуляев Юрий Васильевич**

Сопредседатели: академик **Васильев С.Н.**, член-корр. **Зубарев Ю.Б.**, академик **Кузнецов Н.А.**, академик **Федоров И.Б.**

Члены оргкомитета: проф. Аджемов А.С. (Россия), член-корр. Бабаян Б.А. (Россия), д.т.н. Борисов Ю.И. (Россия), академик Бугаев А.С. (Россия), проф. Вилкова Н.Н. (Россия), проф. Качорек Т. (Польша), проф. Кривошеев М.И. (Россия), проф. Лукин В.В. (Украина), проф. Митра М. (США), проф. Петровский А.А. (Белоруссия), проф. Сарамаки Т. (Финляндия), член-корр. Черепенин В.А. (Россия).

Председатель Национального комитета: член-корр. **Зубарев Юрий Борисович**

Сопредседатели: проф. **Витязев В. В.**, **Кирпичников А.П.**, доц. **Самсонов Г.А.**

Члены оргкомитета: проф. Алпатов Б.А., проф. Бартенев В.Г., проф. Бернюков А.К., проф. Брюханов Ю.А., проф. Васильев К.К., проф. Галушкин А.И., проф. Дворкович В.П., д.т.н. Дворкович А.В., д.т.н. Джиган В.И., Зубарев Р.В., проф. Немировский М.С., проф. Олейников А.Я., проф. Пожидаев В.Н., д.т.н. Приоров А. Л., проф. Прохоров Ю.Н., проф. Санников В.Г., проф. Скачко Ю.В., проф. Сергиенко А.Б., проф. Смольский С.М., проф. Солонина А.И., проф. Соколов И.А., проф. Сперанский В.С., доц. Стешенко В.Б., проф. Сухарев Е.М., проф. Чобану М.К., проф. Шинаков Ю.С., проф. Ямпурин Н.П.

Организация работы в форме пленарных и проблемно-тематических заседаний и заседаний «круглого стола». Пленарные доклады будут представлены по основным направлениям работы Конференции.

Рабочие языки конференции - *русский и английский*. Доклады, включенные в Программу конференции, будут опубликованы на этих же языках в **Сборниках докладов (Трудах) Конференции** к началу ее работы.

РАБОТА КОНФЕРЕНЦИИ ПЛАНИРУЕТСЯ ПО 10 СЕКЦИЯМ:

(укажите соответствующей № секции в заявке к докладу)

1. Теория сигналов и систем
2. Теория и методы ЦОС
3. Обработка сигналов в системах телекоммуникаций
4. Обработка сигналов в радиотехнических системах
5. Обработка и передача изображений
6. Обработка и передача измерительной информации
7. Проектирование и техническая реализация систем ЦОС
8. Цифровое телерадиовещание
9. Цифровая обработка многомерных сигналов
10. Нейрокомпьютерная обработка сигналов и изображений

СРОКИ ПРЕДСТАВЛЕНИЯ ДОКЛАДОВ:

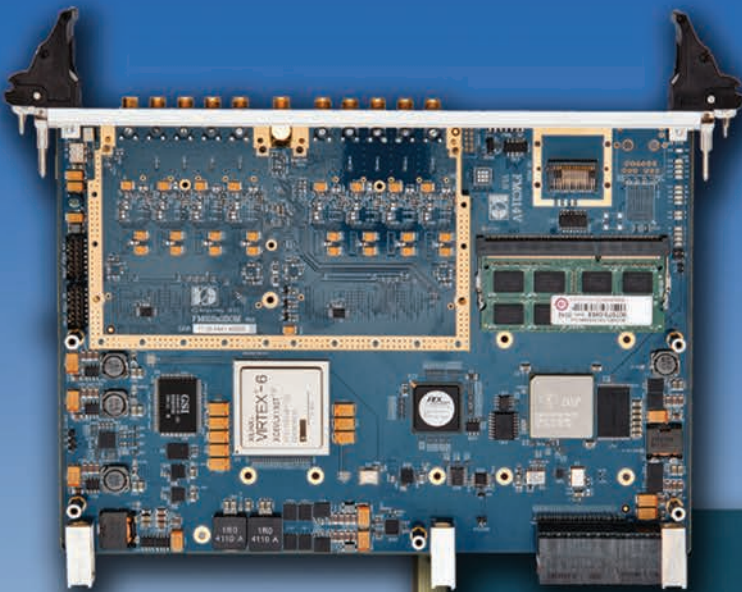
(информация о регистрации обновляется на сайте: <http://www.rntores.ru>)

Реквизиты для перечисления взносов: РНТОРЭС имени А.С. Попова, **ИНН 7702021967, КПП 770201001, БИК 044525225, Р/сч. 40703810038090105080** Московский банк Сбербанка России ОАО, г. Москва.

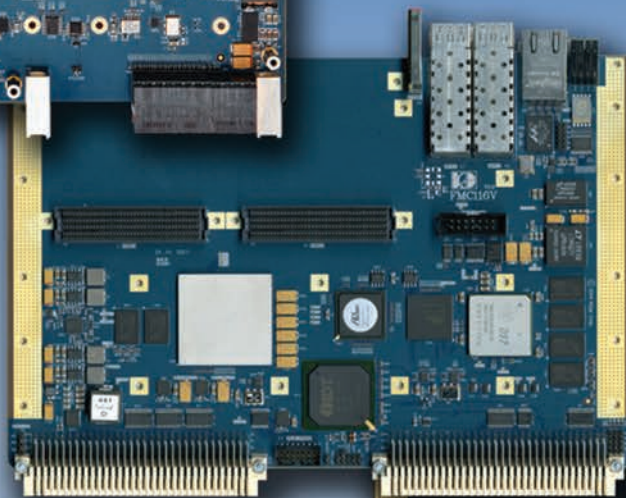
Кор/счет 30101810400000000225

Назначение платежа: «Целевой взнос на конференцию ДСПА-2016», НДС не облагается.

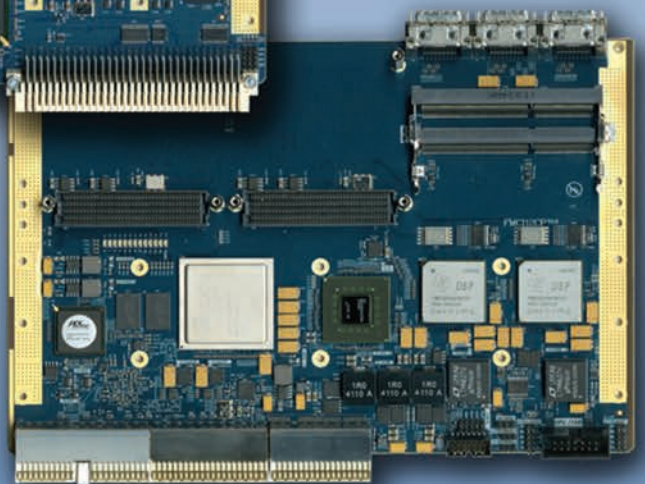
Новый уровень производительности модулей ЦОС



Высокопроизводительные
многоядерные процессоры ЦОС
Масштабируемая модульная
многопроцессорная архитектура



Мезонинные модули
ввода-вывода в стандарте FMC
Программируемая логика ПЛИС
PCI Express, cPCI, VPX, VME
Ethernet, Serial Rapid IO, USB 3.0



Закрытое акционерное общество

“Инструментальные Системы”

www.insys.ru, info@insys.ru