

## МЕТОД КОМПЛЕКСИРОВАНИЯ МНОГОСПЕКТРАЛЬНЫХ ИЗОБРАЖЕНИЙ НА ОСНОВЕ ПЕРЕНОСА ГРАДИЕНТОВ

*Шипко В.В., к.т.н., ВУНЦ ВВС «Военно-воздушная академия им. проф. Н.Е. Жуковского и Ю.А. Гагарина», г. Воронеж, e-mail: shipko.v@bk.ru.*

### METHOD OF INTEGRATION OF MULTISPECTRAL IMAGES ON THE BASIS OF TRANSFER OF GRADIENTS

*Shipko V.V.*

*The new method of integration of digital multispectral images consisting in assignment to one (priority) spectral image of mean or median value of gradients in the sliding window of set of other spectral images is offered. Such approach allows to increase local contrast of the resulting image containing elements of the source images of the same scene received in the different spectral ranges at the same time as much as possible having saved brightness portrait of objects from all spectral a component of the multispectral image and the brightness portrait priority spectral components. The examples of integration images and results of numerical researches confirming efficiency of the offered method are presented.*

**Key words:** integration of images, multispectral images, gradient.

**Ключевые слова:** комплексирование изображений, многоспектральные изображения, градиент.

#### Введение

В настоящее время интенсивное развитие комплексов дистанционного зондирования Земли в основном направлено на расширение спектрального диапазона съемки и увеличение спектрального разрешения. Создаются многоканальные комплексы мониторинга Земной поверхности воздушного и космического базирования [1, 2]. При этом выдвигаются достаточно жесткие требования и на высокое пространственное разрешение таких комплексов. Каждый из приемных каналов многоканальных комплексов имеет свои преимущества и ограничения в информативности получаемых спектральнональных изображений. Это связано со спецификой формирования сигналов в этих спектральных диапазонах, обусловленной различными факторами (спектральная селективность объектов съемки, наличие помех и искажений в отдельных спектральнональных компонентах и т.д.).

Важным направлением обработки получаемых многоспектральных изображений является их комплексирование [3]. Как правило комплексирование заключается в совместной обработке нескольких изображений одного и того же участка местности, полученных в различных спектральных диапазонах, в результате чего можно дополнительно значительно повысить заметность объекта или его деталей на едином изображении. Это происходит благодаря возможности извлечения посредством такой обработки дополнительной информации об объектах местности, заключенной в характере взаимосвязи зональных яркостей между спектральными диапазонами. Конечным продуктом комплексирования может быть как единое полутонное, так и псевдоцветное изображение, на котором с требуемой точностью отображают-

*Предлагается новый метод комплексирования цифровых многоспектральных изображений, заключающийся в присвоении одному (приоритетному) спектральнональному изображению среднего или медианного значения градиентов в скользящем окне совокупности других спектральнональных изображений. Такой подход позволяет повысить локальный контраст результирующего изображения, содержащего элементы исходных изображений одной и той же сцены, полученных в разных спектральных диапазонах, при этом максимально сохранить контурные признаки объектов от всех спектральных компонент многоспектрального изображения и яркостный портрет приоритетной спектральнональной компоненты. Представлены примеры комплексированных изображений и результаты численных исследований, подтверждающие эффективность предложенного метода.*

ся спектральные признаки объектов интереса.

Достоинствами синтеза единого (комплексированного) изображения является:

- возможность аккумулирования в единое изображение особенностей спектральнональных изображений объектов местности;

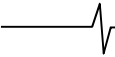
- возможность сделать доступной информацию многоспектральных изображений, связанную с корреляцией яркостей между каналами съемки, а точнее с различиями в корреляциях яркостей объектов и фонов.

Условно, синтез единого полутонного комплексированного изображения по набору спектральнональных изображений можно разделить на два подхода [4]:

- 1) синтез единого полутонного изображения на основе линейной комбинации компонент вектора зональных яркостей;

- 2) синтез единого полутонного изображения на основе нелинейных функций компонент вектора зональных яркостей.

Основной задачей комплексирования является перенос информации составляющих разносспектральных изображений, при этом требуется обеспечить высокую скорость обработки в режиме реального времени с мак-



симальным сохранением информации от всех спектральных составляющих.

Существует множество методов и алгоритмов комплексирования [3, 5-9]. Но большая часть из них эффективна лишь для двухспектральных изображений, либо эти алгоритмы являются вычислительно затратными для многоспектральных изображений.

Как показал анализ [3-9], очень часто при комплексировании многоспектральных изображений выбирают приоритетную спектральную компоненту с целью максимизировать ее влияние на результирующее единое изображение. Поскольку при наличии большого набора спектральных компонент с различными фоно-объектовыми портретами, а часто это инверсные области [3, 10], возникают трудности их качественного объединения, в частности при усреднении таких областей они сольются в один тон. Поэтому выбирается приоритетная спектральная компонента, к которой по определенному правилу добавляются отличительные признаки других компонент. Как известно, наиболее информативными отличительными признаками являются контура объектов [3, 4, 10, 11], поэтому этот факт будет использоваться при постановке задачи синтеза единого комплексированного изображения и оценке его эффективности.

#### Постановка задачи

Используемая модель исходного оцифрованного по строкам  $i$  и столбцам  $j$   $L$ -компонентного изображения  $\Lambda$  в общем случае имеет вид:

$$\Lambda = [\lambda_{i,j}^1 \ \lambda_{i,j}^2 \ \dots \ \lambda_{i,j}^L]^T, \quad \lambda_{i,j}^l \in [0, \dots, 2^N - 1], \quad (1)$$

где  $i = 1, \dots, m, j = 1, \dots, n$ ;  $m, n$  – число строк и столбцов изображения соответственно;  $\lambda_{i,j}^l$  – элементы компоненты  $l$ ;  $l$  – индекс компоненты,  $l = 1, \dots, L, N$  – степень квантования яркости компонент изображения  $\Lambda$ ;

Формально запись процедуры синтеза можно представить следующим образом:

$$\hat{\lambda}_{i,j} = F[\Lambda], \quad (2)$$

где  $\hat{\lambda}_{i,j}$  – скалярная величина, яркость точки  $(i, j)$  единого изображения;  $F$  – некоторое преобразование компонент вектора яркости многоспектрального изображения  $\Lambda$  для точки  $(i, j)$ .

С учетом выбора приоритетной спектральной компоненты выражение (2) можно переписать следующим образом:

$$\hat{\lambda}_{i,j}^r = F[\Lambda], \quad (3)$$

где  $r$  – индекс приоритетной компоненты,  $r \in [1, \dots, L]$ ;

$\hat{\lambda}_{i,j}^r$  – комплексированное изображение с приоритетом  $r$  компоненты.

Таким образом, необходимо синтезировать единое изображение  $\hat{\lambda}_{i,j}^r$  по совокупности спектральных изображений  $\lambda_{i,j}^1, \lambda_{i,j}^2, \dots, \lambda_{i,j}^L$ , которое будет максимально приближено к зональной яркости приоритетного изображения  $\lambda_{i,j}^l$  и содержать контурные признаки

остальных спектральных изображений, т.е. удовлетворять равенству:

$$\hat{\lambda}_{i,j}^r = F1[F2[\Lambda]] \rightarrow \min \left\{ \left\| \hat{\lambda}_{i,j}^r - \lambda_{i,j}^r \right\|, \left\| \hat{g}_{i,j}^r - g_{i,j}^y \right\| \right\}, \quad (4)$$

где  $F1$  – оператор яркостного преобразования;  $F2$  – оператор пространственного (контурного) преобразования;  $\hat{g}_{i,j}^r$  – бинарное изображение пространственного положения контуров единого комплексированного изображения;  $g_{i,j}^y$  – бинарное изображение пространственного положения эталонных контуров, например, определяемое по усредненному изображению всех спектральных компонент  $y_{i,j} = \frac{1}{L} \sum_1^L \lambda_{i,j}^l$ .

#### Метод комплексирования мультиспектральных изображений на основе переноса градиентов спектральных компонент

Вначале определяют приоритетное изображение  $\hat{\lambda}_{i,j}^r$  путем субъективной оценки изображений оператором, либо используя какой-нибудь автоматический критерий.

На следующем шаге формируют единое (эталонное) изображение для оценки общих контурных признаков:

$$y_{i,j} = Y[\lambda_{i,j}^1, \lambda_{i,j}^2, \dots, \lambda_{i,j}^L], \quad (5)$$

где  $Y$  – оператор формирования единого изображения для оценки общих контуров.

В качестве  $Y$  может использоваться любой способ преобразования зональных яркостей, к примеру, усреднение, максимум, линейная комбинация компонент  $\Lambda$  с весовыми коэффициентами и т.д.

После чего вычисляют разности значений яркости (градиенты) каждого пиксела изображения  $y_{i,j}$  с окружающими его пикселями в скользящем окне:

$$(\Delta_{i \pm p, j \pm q})_d = (y_{i,j} - y_{i \pm p, j \pm q})k, \quad (6)$$

где  $p = 0, \dots, P$  – параметр, определяющий координаты пиксела окрестности по  $i$ ;  $q = 0, \dots, Q$  – параметр, определяющий координаты пиксела окрестности по  $j$ ;  $d = 1, \dots, D-1$  – индекс разности значений яркости  $i, j$ -го пиксела с пикселями окрестности;  $D$  – количество пикселей в скользящем окне,  $D = (2P+1)(2Q+1)$ ;  $k$  – коэффициент усиления градиентов.

Далее формируют набор оценок каждого пиксела единого комплексированного изображения с приоритетной компонентой  $r$  по полученным градиентам единого эталонного изображения. Это может быть выполнено в виде суммы набора градиентов (6) и каждого пиксела приоритетного изображения  $\lambda_{i,j}^r$  [9]:

$$(\hat{\lambda}_{i,j}^r)_d = \lambda_{i,j}^r + (\Delta_{i \pm p, j \pm q})_d. \quad (7)$$

Результирующее единое комплексированное изображение может быть представлено в виде среднего или медианного значения оценок  $(\hat{\lambda}_{i,j}^r)_d$ :

$$\hat{\lambda}_{i,j}^r = \text{MEAN}_d[(\hat{\lambda}_{i,j}^r)_d] = \frac{1}{D-1} \sum_d (\hat{\lambda}_{i,j}^r)_d, \quad (8)$$

$$\hat{\lambda}_{i,j}^r = \text{MED}[(\hat{\lambda}_{i,j}^r)_d], \quad (9)$$

Таким образом, при комплексировании многоспектральных изображений по данному методу происходит перенос градиентов всех спектрозональных составляющих наиболее приоритетному в яркостном смысле изображению.

**Примеры комплексирования и результаты численных исследований**

В качестве критерия различия яркостей приоритетного изображения  $\lambda_{i,j}^r$  и комплексированного изображения  $\hat{\lambda}_{i,j}^r$ , выбрана среднеквадратическая ошибка (СКО):

$$\sigma = \sqrt{\frac{1}{mn} \sum_{i=1}^m \sum_{j=1}^n (\hat{\lambda}_{i,j}^r - \lambda_{i,j}^r)^2}. \quad (10)$$

Для исследования качественных характеристик предлагаемого метода в сохранении контурных признаков всех спектрозональных компонент в комплексированном изображении, вычислялись ошибки типа пропуск контура ( $\delta_{\text{ПК}}$ ) и ложное определение контура ( $\delta_{\text{ЛК}}$ ) в соответствии с выражениями

$$\delta_{\text{ЛК}} = \frac{1}{mn} \sum_{i=1}^m \sum_{j=1}^n \begin{cases} 1, & (\hat{g}_{i,j}^r - g_{i,j}^y) = 1, \\ 0, & (\hat{g}_{i,j}^r - g_{i,j}^y) \neq 1, \end{cases} \quad (11)$$

$$\delta_{\text{ПК}} = \frac{1}{mn} \sum_{i=1}^m \sum_{j=1}^n \begin{cases} 1, & (g_{i,j}^y - \hat{g}_{i,j}^r) = 1, \\ 0, & (g_{i,j}^y - \hat{g}_{i,j}^r) \neq 1, \end{cases} \quad (12)$$

при этом общая (суммарная) ошибка определения контуров

$$\delta = \delta_{\text{ПК}} + \delta_{\text{ЛК}}. \quad (13)$$

На рис. 1 представлены шесть исходных спектрозональных изображения, а на рис. 2 – единые эталонные изображения для оценки контуров, полученные методом

усреднения исходных изображений  $y_{i,j}^{\text{mean}} = \frac{1}{L} \sum_l \lambda_{i,j}^l$  (а),

методом максимума  $y_{i,j}^{\text{max}} = \max_l [\lambda_{i,j}^l]$  (б) и среднего между ними

$y_{i,j}^{\text{max mean}} = \frac{(y_{i,j}^{\text{mean}} + y_{i,j}^{\text{max}})}{2}$  (в). На рис. 3 представлены

бинарные контурные изображения, полученные детектором границ Канны [11] с фиксированными параметрами. Из рис. 3 видно, что на контурах усредненного изображения  $y_{i,j}^{\text{mean}}$  хорошо просматриваются основные объекты, за исключением некоторых объектов (правый верхний угол изображений  $\lambda_{i,j}^l$ ), при этом на контурах изображения  $y_{i,j}^{\text{max}}$  эти участки хорошо прорисованы. Альтернативой является изображение  $y_{i,j}^{\text{max mean}}$ , на контурах которого хорошо отображены детали объектов изображений  $y_{i,j}^{\text{mean}}$  и  $y_{i,j}^{\text{max}}$ .

В табл. 1 для сравнения представлены значения взаимосвязей (10) и (13) между спектрозональными изображениями  $\lambda_{i,j}^l$  и эталонами  $y_{i,j}^{\text{mean}}$ ,  $y_{i,j}^{\text{max}}$ ,  $y_{i,j}^{\text{max mean}}$ .

На рис. 4, 5 представлены результаты численных исследований предлагаемого метода комплексирования. При проведении вычислительного эксперимента использовались спектрозональные изображения, представленные на рис. 1. В качестве приоритетного изображения при комплексировании и оценке яркости выбрано первое спектрозональное изображение  $\lambda_{i,j}^1$ , в качестве эталонов для переноса градиентов и последующей оценки контуров

использовались изображения  $y_{i,j}^{\text{mean}}$ ,  $y_{i,j}^{\text{max}}$ ,  $y_{i,j}^{\text{max mean}}$ , контура которых представлены на рис. 3, конечная оценка вычислялась по выражению (9).

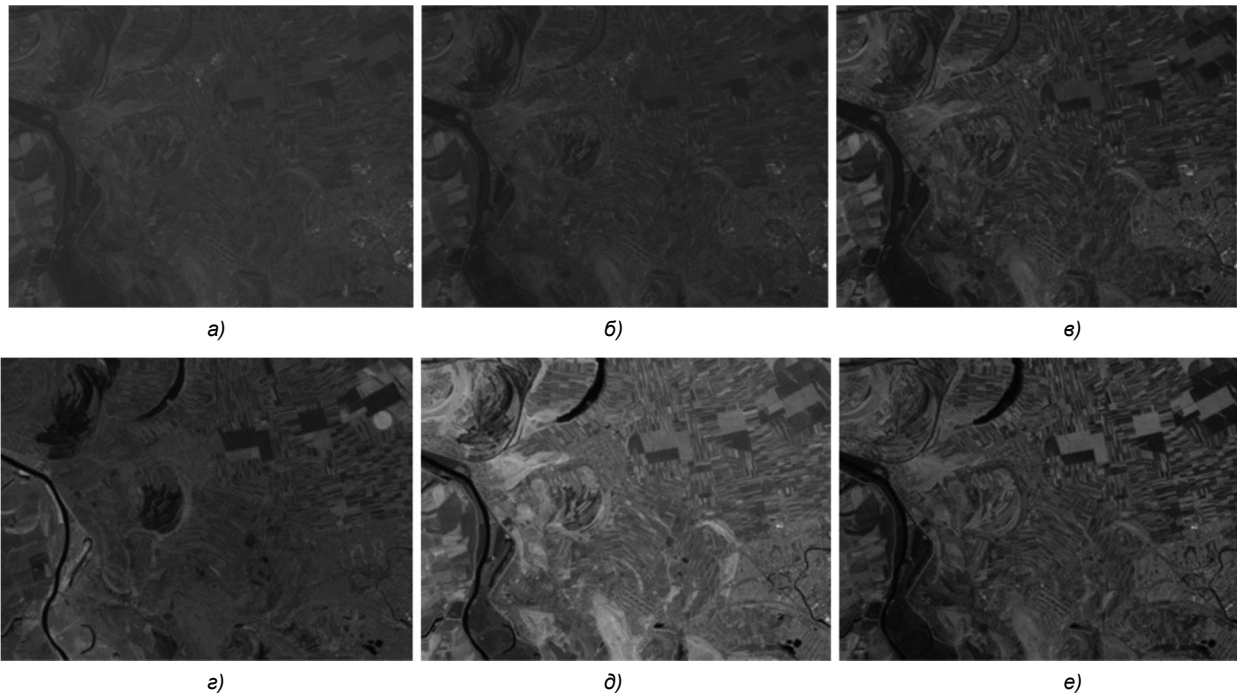


Рис. 1. Спектрозональные изображения:  $\lambda_{i,j}^1$  (480 нм) – а,  $\lambda_{i,j}^2$  (550 нм) – б,  $\lambda_{i,j}^3$  (670 нм) – в,  $\lambda_{i,j}^4$  (820 нм) – г,  $\lambda_{i,j}^5$  (1650 нм) – д,  $\lambda_{i,j}^6$  (2200 нм) – е

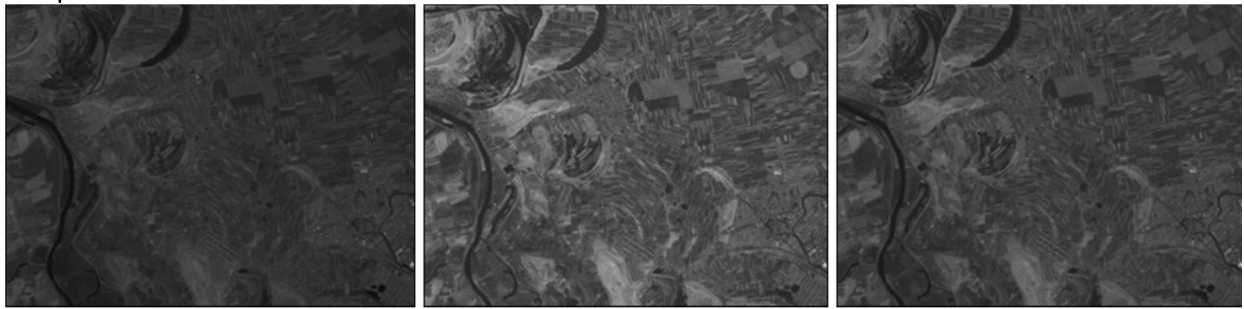


Рис. 2. Эталонные комплексированные изображения:  $y_{i,j}^{\text{mean}}$  – а,  $y_{i,j}^{\text{max}}$  – б,  $y_{i,j}^{\text{max, mean}}$  – в

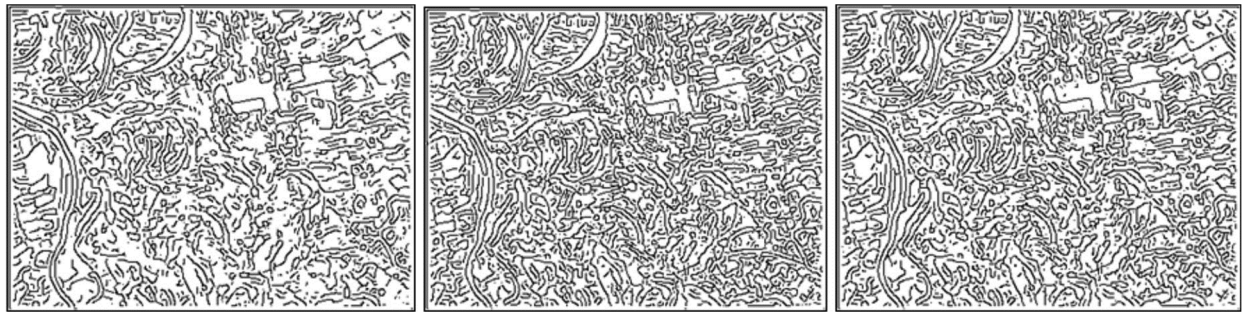


Рис. 3. Контура изображений:  $y_{i,j}^{\text{mean}}$  – а,  $y_{i,j}^{\text{max}}$  – б,  $y_{i,j}^{\text{max, mean}}$  – в

Таблица 1 – Яркостные и контурные взаимосвязи исходных изображений и эталонов

	$\lambda_{i,j}^1$	$\lambda_{i,j}^2$	$\lambda_{i,j}^3$	$\lambda_{i,j}^4$	$\lambda_{i,j}^5$	$\lambda_{i,j}^6$
$y_{i,j}^{\text{mean}}$	$\sigma = 7,70$ $\delta = 0,138$	$\sigma = 35,50$ $\delta = 0,136$	$\sigma = 11,45$ $\delta = 0,126$	$\sigma = 21,57$ $\delta = 0,241$	$\sigma = 33,63$ $\delta = 0,123$	$\sigma = 13,89$ $\delta = 0,162$
$y_{i,j}^{\text{max}}$	$\sigma = 35,48$ $\delta = 0,203$	$\sigma = 47,56$ $\delta = 0,209$	$\sigma = 44,60$ $\delta = 0,189$	$\sigma = 49,23$ $\delta = 0,261$	$\sigma = 11,01$ $\delta = 0,057$	$\sigma = 40,61$ $\delta = 0,188$
$y_{i,j}^{\text{max, mean}}$	$\sigma = 18,54$ $\delta = 0,179$	$\sigma = 29,96$ $\delta = 0,184$	$\sigma = 27,29$ $\delta = 0,160$	$\sigma = 33,61$ $\delta = 0,254$	$\sigma = 17,63$ $\delta = 0,08$	$\sigma = 24,62$ $\delta = 0,171$

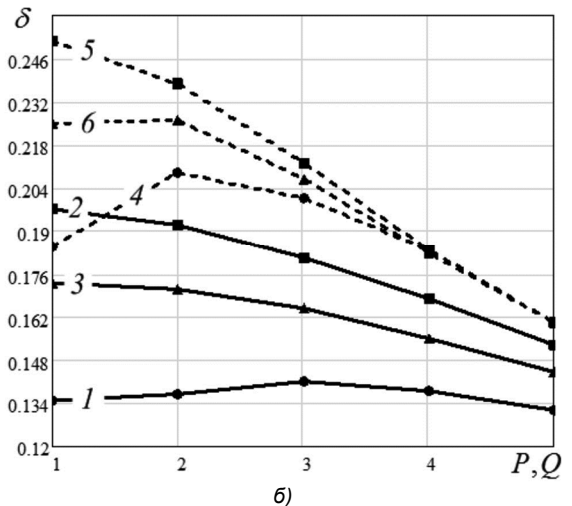
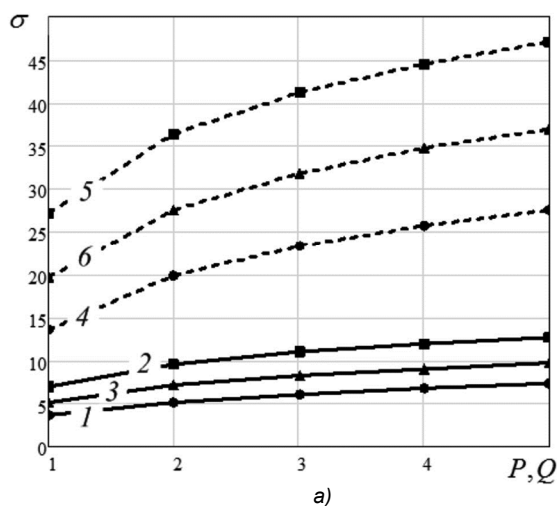


Рис. 4. Зависимость  $\sigma$  (а) и  $\delta$  (б) от размера скользящего окна  $(P, Q)$ : 1 –  $y_{i,j}^{\text{mean}}$ ,  $k=1$ , 2 –  $y_{i,j}^{\text{max}}$ ,  $k=1$ , 3 –  $y_{i,j}^{\text{max, mean}}$ ,  $k=1$ , 4 –  $y_{i,j}^{\text{mean}}$ ,  $k=4$ , 5 –  $y_{i,j}^{\text{max}}$ ,  $k=4$ , 6 –  $y_{i,j}^{\text{max, mean}}$ ,  $k=4$

На рис. 6-9 представлены некоторые варианты комплексированных изображений, полученные по раз-

работанному методу с результатами их численных оценок.



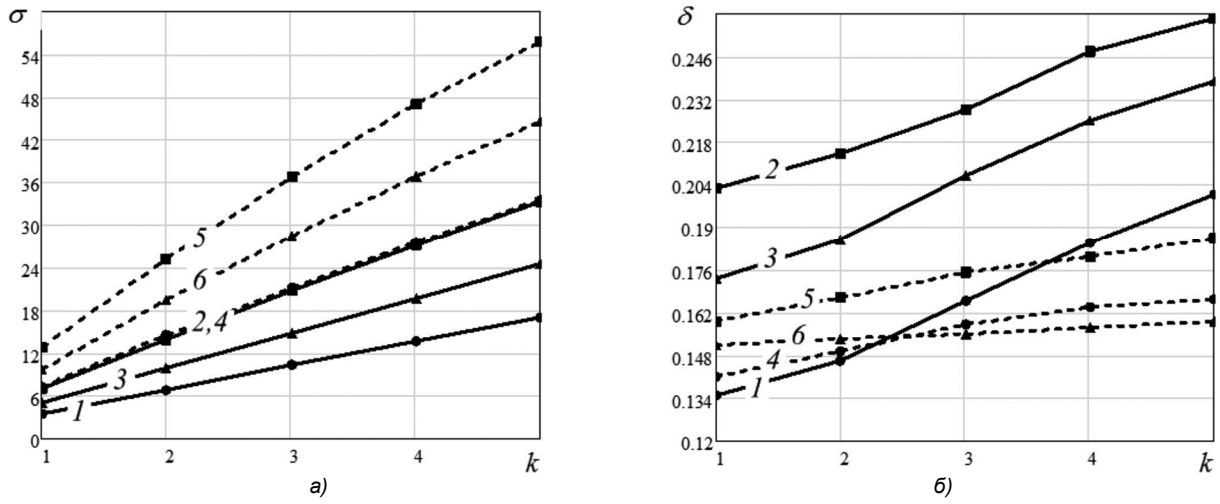
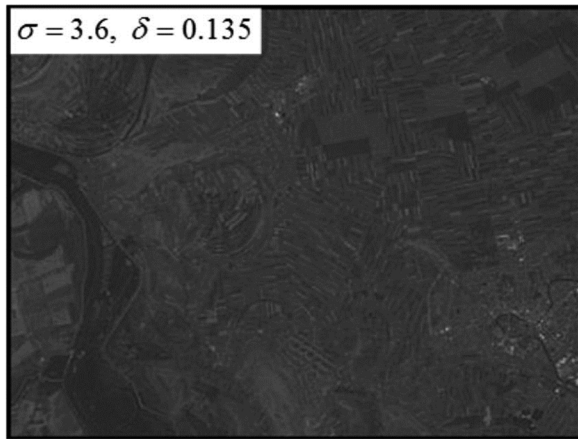
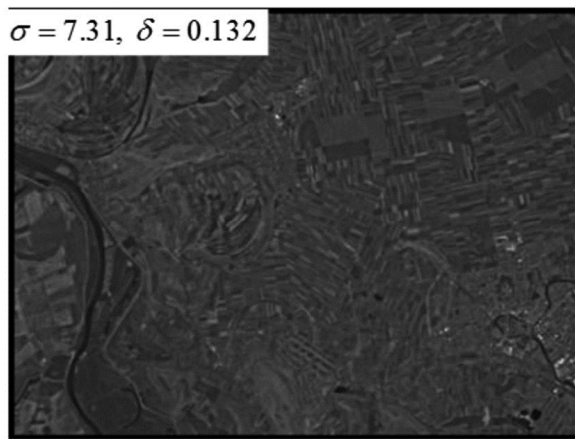


Рис. 5. Зависимость  $\sigma$  (а) и  $\delta$  (б) от коэффициента усиления градиентов  $k$  :

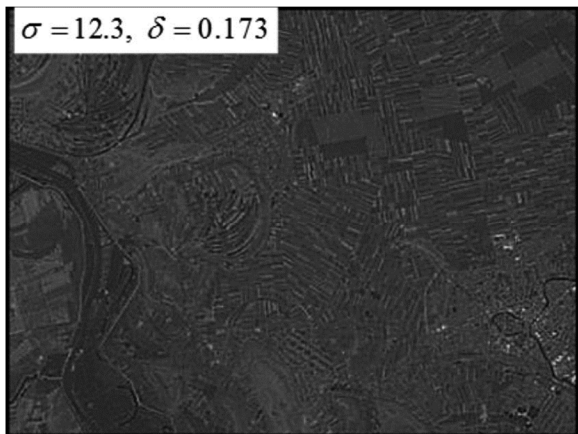
$1 - y_{i,j}^{mean}, (P,Q)=1, 2 - y_{i,j}^{max}, (P,Q)=1,$   
 $3 - y_{i,j}^{max, mean}, (P,Q)=1, 4 - y_{i,j}^{mean}, (P,Q)=5, 5 - y_{i,j}^{max}, (P,Q)=5, 6 - y_{i,j}^{max, mean}, (P,Q)=5$



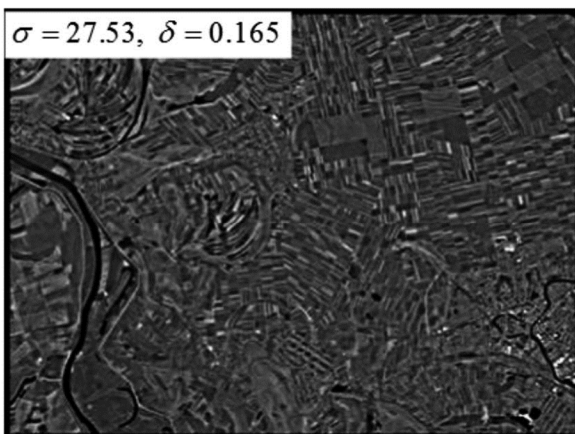
а)



б)

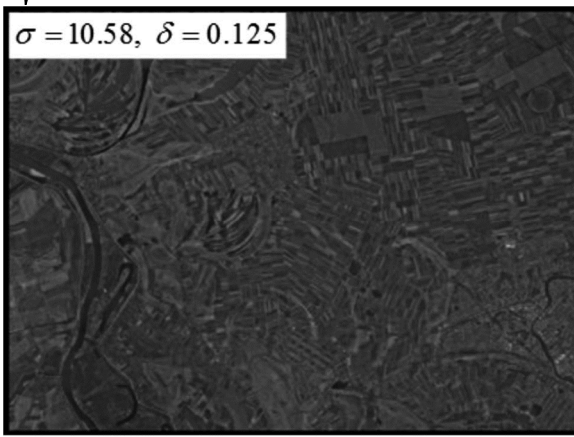


в)

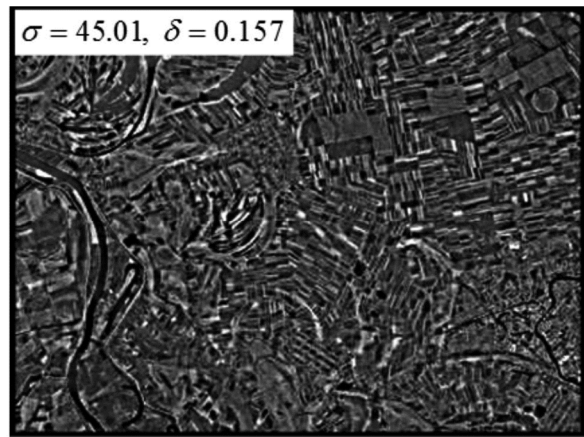


г)

Рис. 6. Комплексированные изображения, полученные по выражению (9) с приоритетом яркости изображения  $\lambda_{i,j}^1$  и переносом градиентов изображения  $y_{i,j}^{mean}$  :  
 $(P,Q)=1, k=1, - а, (P,Q)=5, k=1 - б, (P,Q)=1, k=4 - в, (P,Q)=5, k=4 - г$

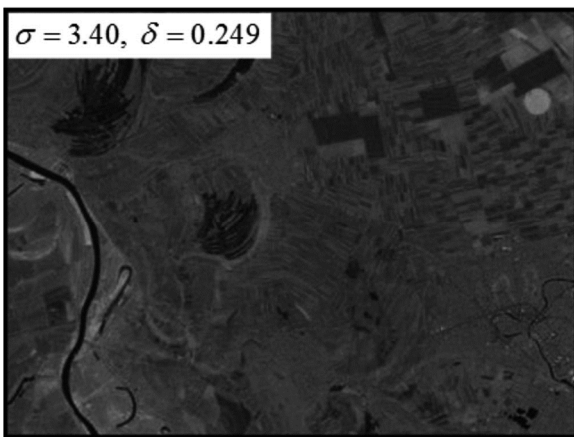


а)

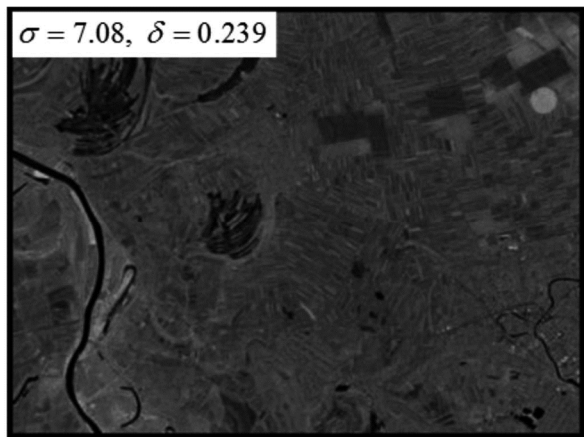


б)

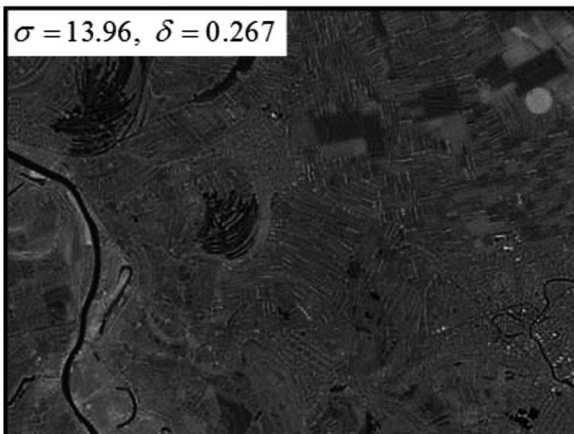
Рис. 7. Комплексированные изображения, полученные по выражениям (7), (9) с приоритетом яркости изображения  $\lambda_{i,j}^1$  и переносом градиентов изображения  $y_{i,j}^{\max}$ :  $(P,Q)=5, k=1$  – а,  $(P,Q)=1, k=4$  – б



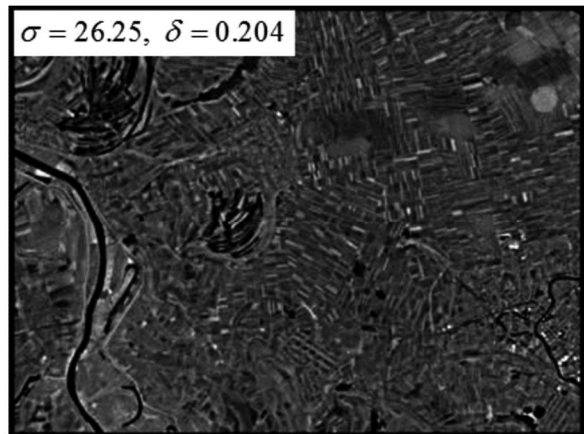
а)



б)



в)



г)

Рис. 8. Комплексированные изображения, полученные по выражениям (7), (9) с приоритетом яркости изображения  $\lambda_{i,j}^4$  и переносом градиентов изображения  $y_{i,j}^{\max}$ :  $(P,Q)=1, k=1$  – а,  $(P,Q)=5, k=1$  – б,  $(P,Q)=1, k=4$  – в,  $(P,Q)=5, k=4$  – г

Анализ результатов проведенных исследований показал, что результирующие комплексированные изображения обладают высоким локальным контрастом, яркостью близкой к приоритетному изображению и с контурными признаками всех спектральных составляющих. Выявлено, что с увеличением коэффициента усиления  $k$  можно повысить локальный контраст, что является положительным визуальным эффектом, но при этом увеличи-

вается СКО (10) с приоритетным изображением, а также могут возникать ложные контурные признаки, обусловленные шумовой составляющей, что приводит к увеличению (13). С увеличением размеров скользящего окна, также улучшается визуальное качество изображения и критерий контурных признаков (13), при этом возможно небольшое возрастание СКО (10) и вычислительных затрат алгоритма.

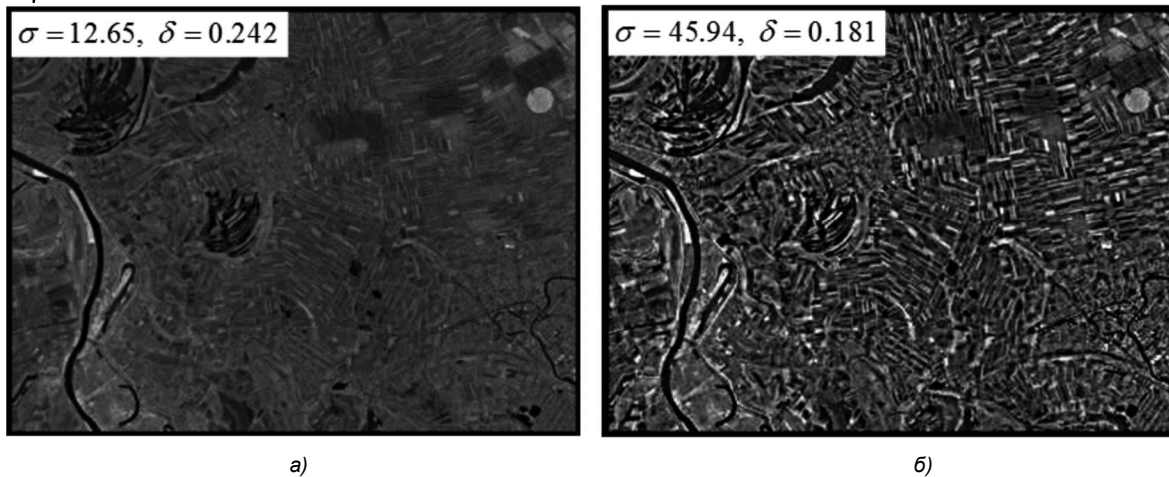


Рис. 9. Комплексированные изображения, полученные по выражениям (7), (9) с приоритетом яркости изображения  $\lambda_{i,j}^4$  и переносом градиентов изображения  $y_{i,j}^{\max}$  :

$$(P, Q) = 5, k = 1 - а, (P, Q) = 5, k = 4 - б$$

### Заключение

Таким образом, предлагаемый метод позволяет с высокой эффективностью комплексировать многоспектральные изображения. Результирующее комплексированное изображение обладает яркостными признаками приоритетного спектрального изображения и контурными признаками всех остальных спектральных изображений. Разработанные алгоритмы на основе предлагаемого метода могут использоваться в автоматизированных комплексах сбора и обработки данных дистанционного зондирования Земли с целью улучшения качества комплексирования.

### Литература

1. Тарасов В.В., Якушенков Ю.Г. Двух- и многодиапазонные оптико-электронные системы с матричными приемниками излучения. М.: Университетская книга; Логос. 2007. 192 с.
2. Действующие и перспективные спутники наблюдения земли // Геоматика. 2013. № 2. С. 107-111.
3. Современные технологии обработки данных дистанционного зондирования Земли / Под ред. В.В. Еремеева. М.: Физматлит, 2015. 460 с.
4. Автоматизированные системы наземных комплек-

сов сбора и обработки данных воздушной разведки. Учебное пособие / Под ред. И.Н. Белоглазова. М.: ВВИА им. проф. Н.Е. Жуковского. 2003. 296 с.

5. Патент RU 2451338 опубликовано 20.05.2012 г., МПК G06Т 5/00

6. Патент RU 2540778 опубликовано 10.02.2015 г., МПК G06Т 5/50, G06К 9/46

7. Богданов А.П., Холопов И.С. Алгоритмы формирования цветного комплексированного изображения из многоспектральных монохромных на основе методов преобразования цветов / Цифровая обработка сигналов. 2013. № 3. С. 26-32.

8. Васильев А.С. Методы комплексирования изображений многоспектральных оптико-электронных систем // Сборник трудов международной конференции «Прикладная оптика – 2014». – СПб, 2014, Т.2, С.191-194.

9. Шипко В.В. Алгоритм комплексирования двух разноспектральных изображений на основе цветового синтеза // Цифровая обработка сигналов. 2017. № 3. С. 32-38.

10. Сагдуллаев Ю.С., Ковин С.Д. Восприятие и анализ разноспектральных изображений: Монография. М.: Издательство «Спутник+», 2016. 251 с.

11. Гонсалес Р., Вудс Р. Цифровая обработка изображений. М.: Техносфера, 2005. 1072 с.

## О ПОВЫШЕНИИ ПОМЕХОУСТОЙЧИВОСТИ МЕТОДА КОМПЛЕКСИРОВАНИЯ МНОГОСПЕКТРАЛЬНЫХ ИЗОБРАЖЕНИЙ НА ОСНОВЕ ПЕРЕНОСА ГРАДИЕНТОВ

*Шипко В.В., к.т.н., ВУНЦ ВВС «Военно-воздушная академия им. проф. Н.Е. Жуковского и Ю.А. Гагарина», г. Воронеж, e-mail: shipko.v@bk.ru.*

## ABOUT INCREASE OF THE NOISE STABILITY OF THE METHOD OF INTEGRATION OF MULTISPECTRAL IMAGES ON THE BASIS OF TRANSFER OF GRADIENTS

*Shipko V.V.*

*Approach to increase of a noise stability of a method of integration of multispectral images on the basis of transfer of gradients is given in article. It is shown that estimates of brightness priority spectral components can be created by a method of interchannel gradient reconstruction. The presented examples of integration images and results of numerical researches, confirm efficiency of the offered approach.*

**Key words:** integration of images, multispectral images, noise stability.

**Ключевые слова:** комплексирование изображений, многоспектральные изображения, помехоустойчивость.

### Введение

Комплексирование многоспектральных изображений является одним из самых важных направлений обработки изображений. Адекватным способом комплексирования разнородных многоспектральных изображений является комплексирование с приоритетом одной из спектральных компонент с целью максимизировать ее влияние на результирующее единое изображение. При этом необходимо перенести отличительные признаки остальных спектральных изображений на результирующее изображение. Наиболее информативным признаком изображений объектов являются их контуры [1-4]. Поэтому именно контуры изображений всех спектральных компонент многоспектрального изображения должны присутствовать в едином комплексированном изображении.

Хорошие результаты комплексирования с приоритетом, показал метод на основе переноса градиентов [5]. Но качественное выделение и перенос контурных признаков приоритетному изображению могут затруднять помехи и различного рода искажения. Поэтому необходимо учитывать эти обстоятельства при построении алгоритмов комплексирования на основе переноса градиентов.

### Пути повышения помехоустойчивости метода комплексирования мультиспектральных изображений на основе переноса градиентов

Очень важно, при формировании единого (эталонного) изображения для оценки общих контурных признаков выбрать адекватный метод. Например, метод усреднения  $L$ -компонент, частично сгладит шумовую составляющую, однако не всегда это даст хороший результат [2, 3], и такому изображению потребуются дополнительная фильтрация, что может привести к потере ценной информации о контурах объектов. В свою очередь метод максимума

*Приведен подход к повышению помехоустойчивости метода комплексирования многоспектральных изображений на основе переноса градиентов. Показано, что оценки яркости приоритетной спектральной компоненты можно формировать по методу межканальной градиентной реконструкции. Представленные примеры комплексированных изображений и результаты численных исследований, подтверждают эффективность предложенного подхода.*

может наоборот повысить интенсивность шума на едином изображении. Поэтому вопрос формирования эталона контурных признаков является очень важной задачей и требует отдельного рассмотрения.

После получения эталона контуров, как указано в [5], формируется набор оценок каждого пиксела единого комплексированного изображения с приоритетной компонентой  $r$  по полученным градиентам единого эталонного изображения:

$$(\hat{\lambda}_{i,j}^r)_d = \lambda_{i,j}^r + (\Delta_{i\pm p, j\pm q})_d \quad (1)$$

При формировании набора оценок (1), в случае наличия шума или различного рода искажений в приоритетном канале, эти искажения перейдут и в комплексированное изображение.

Хороший результат по устранению как импульсного, так и аддитивного шума на многоспектральных изображениях показал метод межканальной градиентной реконструкции [6, 7], заключающийся в восстановлении искаженных пикселей одной спектральной компоненты по градиентам соседних компонент. В соответствии с методом межканальной градиентной реконструкции, набор оценок комплексированного изображения можно сформировать следующим образом:

$$(\hat{\lambda}_{i,j}^r)_d = \lambda_{i\pm p, j\pm q}^r + (\Delta_{i\pm p, j\pm q})_d \quad (2)$$

Тогда единое комплексированное изображение можно получить в виде оценки:

$$\hat{\lambda}_{i,j}^r = M[(\hat{\lambda}_{i,j}^r)_d], \quad (3)$$

где  $M$  – оператор формирования оценки единого комплексированного изображения [5].

**Примеры комплексирования и результаты численных исследований**

При проведении численных исследований использовались спектрозональные изображения, представленные в [5]. В качестве приоритетного изображения при комплексировании и оценке яркости выбрано первое спектрозональное изображение  $\lambda_{i,j}^1$ , в качестве эталонов для переноса градиентов и последующей оценки контуров использовалось среднее изображение  $y_{i,j}^{mean}$ . Конечная оценка (3) вычислялась как **медиана** набора оценок (1) или (2). На рис. 1 представлены зависимости  $\sigma$  и  $\delta$  [5]

от СКО аддитивного гауссовского шума  $\sigma_{ш}$  при разных значениях  $P, Q$ . На рис. 1 кривая 1 – для оценок (1) при  $(P, Q) = 1$ , 2 – для оценок (1) при  $(P, Q) = 5$ , 3 – для (2) при  $(P, Q) = 1$ , 4 – для (2) при  $(P, Q) = 5$ . Кривые 1-4 получены при наличии шума в каждой спектрозональной компоненте. Кривые 5 и 6 получены для оценки (2) при  $(P, Q) = 1$  и  $(P, Q) = 5$  соответственно, в случае наличия шума в одной (приоритетной) компоненте.

На рис. 2, 3 представлены некоторые варианты комплексированных изображений, полученные по разработанному методу с результатами их численных оценок.

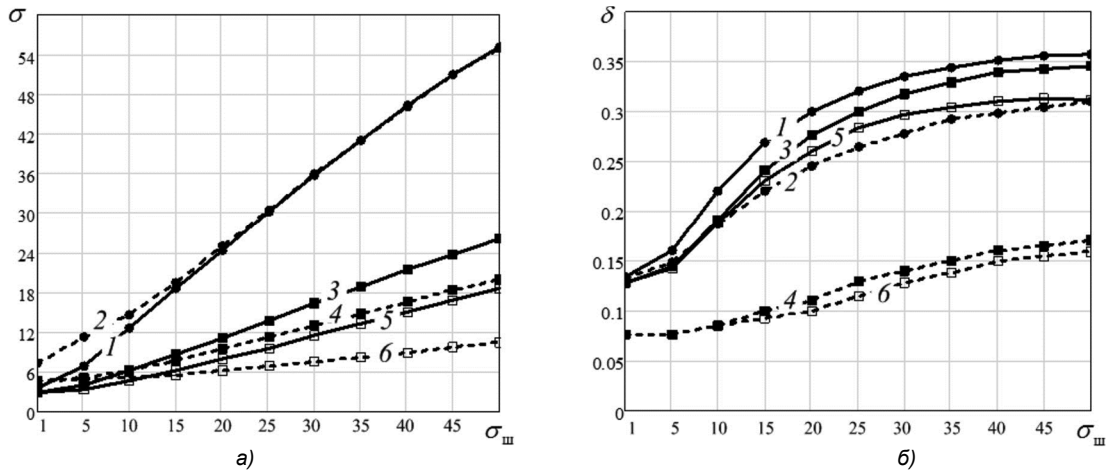


Рис. 1. Зависимость  $\sigma$  (а) и  $\delta$  (б) от СКО шума  $\sigma_{ш}$

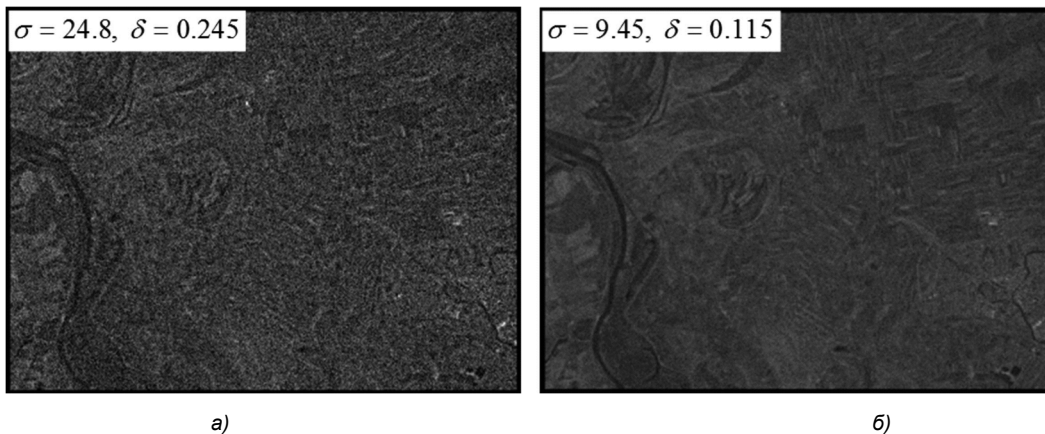


Рис. 2. Комплексированные изображения при наличии шума  $\sigma_{ш} = 20$  в каждом канале, для следующих оценок: (1) при  $(P, Q) = 5$  – а; (2) при  $(P, Q) = 5$  – б

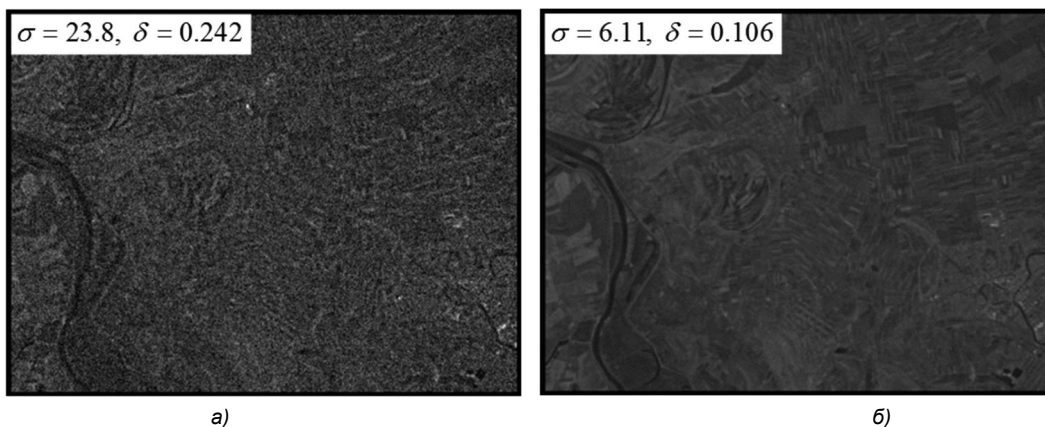
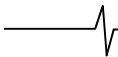


Рис. 3. Комплексированные изображения при наличии шума  $\sigma_{ш} = 20$  в одном канале (приоритетном), для следующих оценок: (1) при  $(P, Q) = 5$  – а; (2) при  $(P, Q) = 5$  – б



Рассмотрим еще один пример. На рис. 4 представлены спектрозональные компоненты гиперспектрального изображения с различного рода искажениями и помехами.

На рис. 5 показаны изображения и их контура, полученные усреднением и максимумом исходных спектрозональных изображений. В качестве эталона для получения контуров в комплексированном изображении по разработанному методу предпочтительней выбрать усредненное изображение  $y_{i,j}^{\text{mean}}$  (рис. 5 а), т.к. на изображении  $y_{i,j}^{\text{max}}$  (рис. 5 б) присутствуют помехи, имеющие максимумы в исходных спектрозональных изображении-

ях, и которые отрицательно повлияют на результирующее изображение.

На рис. 6 представлены комплексированные изображения, полученные по разработанному методу усреднением оценок (2) с приоритетом спектрозональной компоненты  $\lambda_{i,j}^4$  и эталоном для переноса контуров  $y_{i,j}^{\text{mean}}$ . Полученные комплексированные изображения имеют выраженные контурные признаки эталонного изображения  $y_{i,j}^{\text{mean}}$ , при этом яркостный портрет исходного (приоритетного) изображения сохранился. Искажения и помехи присутствующие в исходных спектрозональных изображениях, на результирующем изображении практически отсутствуют.

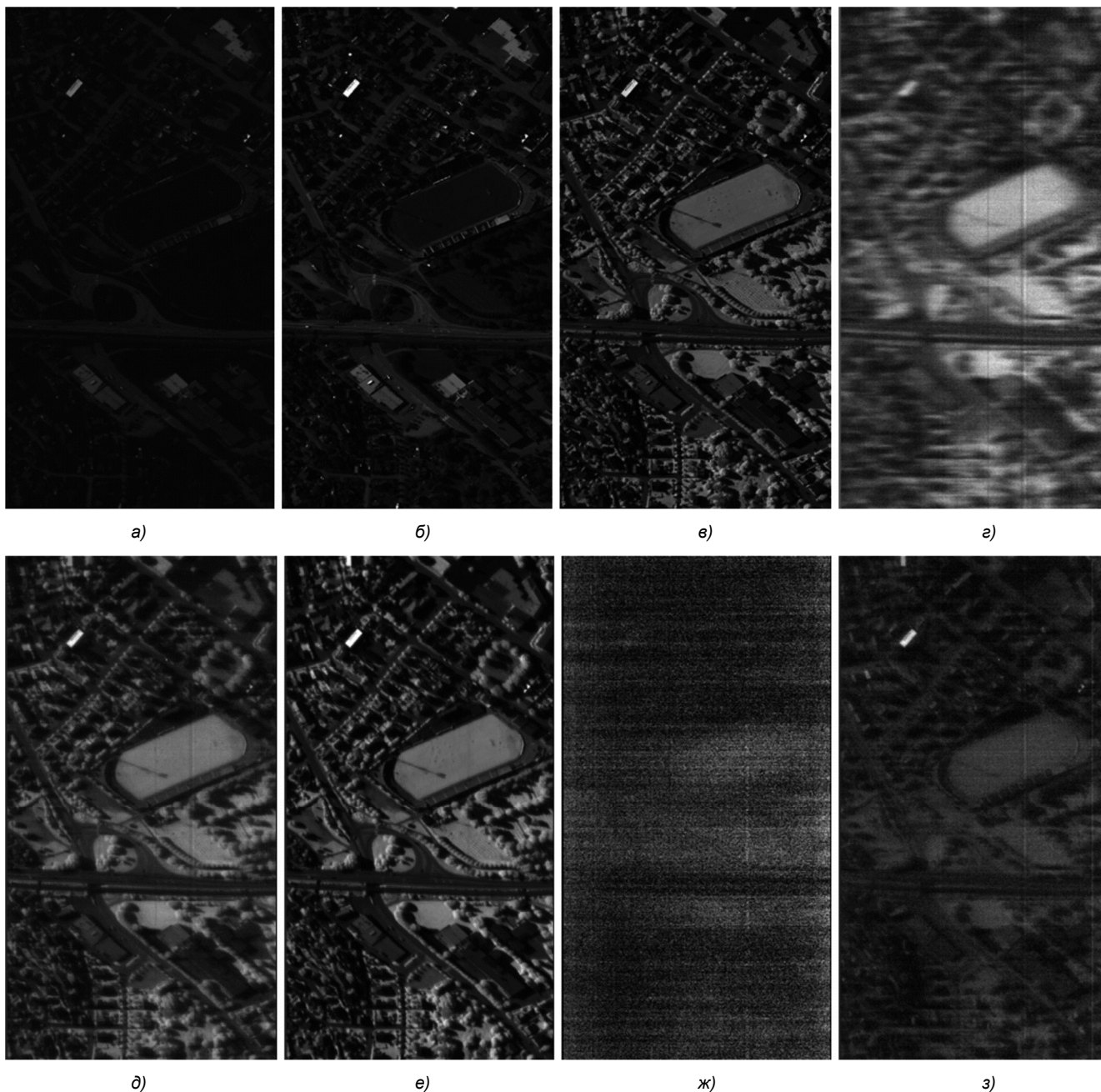


Рис. 4. Спектрозональные изображения:  $\lambda_{i,j}^1$  (401 нм) – а,  $\lambda_{i,j}^2$  (537,8 нм) – б,  $\lambda_{i,j}^3$  (869,8 нм) – в,  $\lambda_{i,j}^4$  (903,4 нм) – г,  $\lambda_{i,j}^5$  (918,6 нм) – д,  $\lambda_{i,j}^6$  (1065,5 нм) – е,  $\lambda_{i,j}^7$  (1364,7 нм) – ж,  $\lambda_{i,j}^8$  (1733,8 нм) – з



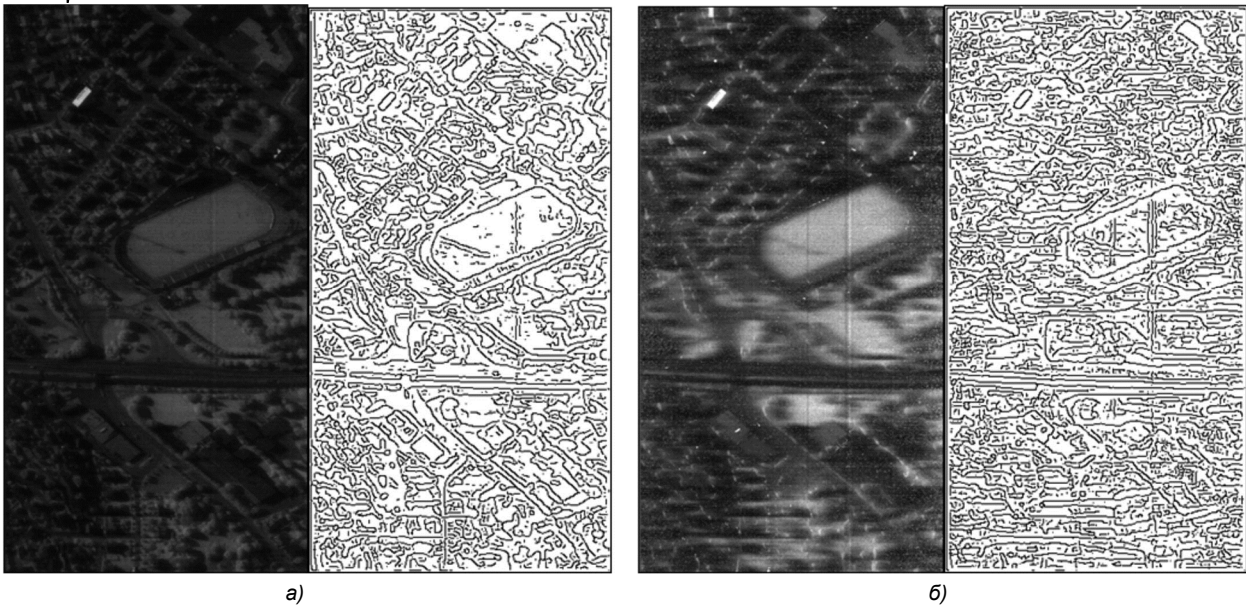


Рис. 5. Эталонные комплексированные изображения и их контура:  $y_{i,j}^{\text{mean}}$  – а,  $y_{i,j}^{\text{max}}$  – б

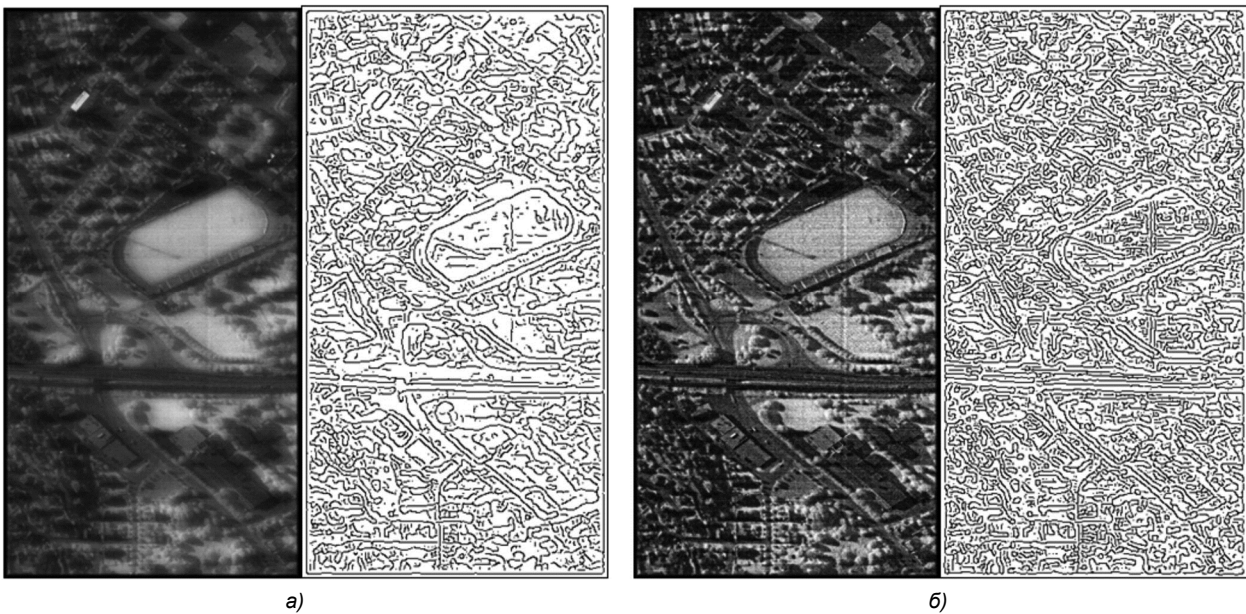


Рис. 6. Комплексированные изображения, полученные усреднением оценок (2) с приоритетом яркости изображения  $\lambda_{i,j}^4$  и переносом градиентов изображения  $y_{i,j}^{\text{mean}}$ :  $(P, Q) = 1, k = 1$ , – а,  $(P, Q) = 5, k = 4$ , – б

## Заключение

Таким образом, разработанный метод позволяет с высокой эффективностью комплексировать многоспектральные изображения, в том числе и гиперспектральные, в условиях помех и различного рода искажений. Дальнейшим направлением развития предложенного метода на основе переноса градиентов является его оптимизация в смысле критериев минимума СКО яркости приоритетного и комплексированного изображений и минимума ошибок определения контуров комплексированного и эталонного изображений.

## Литература

1. Современные технологии обработки данных дистанционного зондирования Земли / Под ред. В.В. Еремеева. М.: Физматлит, 2015. 460 с.
2. Автоматизированные системы наземных комплексов сбора и обработки данных воздушной разведки.

Учебное пособие / Под ред. И.Н. Белоглазова. М.: ВВИА им. проф. Н.Е. Жуковского. 2003. 296 с.

3. Сагдуллаев Ю.С., Ковин С.Д. Восприятие и анализ разноспектральных изображений: Монография. М.: Издательство «Спутник+», 2016. 251 с.

4. Гонсалес Р., Вудс Р. Цифровая обработка изображений. М.: Техносфера, 2005. 1072 с.

5. Шипко В.В. Метод комплексирования многоспектральных изображений на основе переноса градиентов // Цифровая обработка сигналов. 2019. № 3.

6. Самойлин Е.А., Шипко В.В. Метод межканальной градиентной реконструкции искаженных сигналов цветных цифровых изображений // Цифровая обработка сигналов. 2013. № 3. С. 13-16.

7. Самойлин Е.А., Шипко В.В. Межканальная градиентная реконструкция сигналов цветных цифровых изображений искаженных импульсными помехами // Автометрия. 2014. Т. 50. № 2. С. 22-30.



## МЕТОД ВЫСОКОТОЧНОГО КАРТОГРАФИРОВАНИЯ ПО ДАННЫМ СТЕРЕОСЪЁМКИ И ЛАЗЕРНОГО ДАЛЬНОМЕРА

*Кузнецов А.Е., д.т.н., зам. директора НИИ «Фотон» РГРТУ, г. Рязань, e-mail: foton@rsreu.ru.*

### THE HIGH ACCURACY MAPPING METHOD BY STEREO IMAGING DATA

*Kuznetsov A.E.*

*In the paper are described a problems of the Earth surface large scale mapping by remote sensing stereo data without using of ground control points (GCP). It is shown that height detection accuracy by stereo images must be in some times higher than accuracy of plane coordinates of on ground objects. The analysis showed that because of errors of satellite external orientation elements determination the systematical errors are in the on ground objects height detection calculations. Experiment results confirming of systematical errors while making of surface height map are shown in the paper. The algorithm of systematical errors compensation by using measurements of LIDAR is presented. The formula of the method which allows to build a more accurate height map is shown. The orthoplan of mapping territory which is not provided by GCP and digital relief model is build using this height map.*

**Key words:** stereo images, LIDAR, model of stereoprocessing, height map, systematical errors, correction coefficient.

**Ключевые слова:** стереоизображения, лазерный дальномер, модель стереообработки, карта высот, систематические погрешности, поправочный коэффициент.

#### Введение

Стереоизображения предназначены для вычисления трёхмерных координат наблюдаемых территорий и формирования крупномасштабных карт местности, например, масштаба 1:25000. На картах такого масштаба плановые координаты объектов земной поверхности определяются со среднеквадратической ошибкой (СКО) не более 12 м, а их высоты с СКО 2,5-3 м. Для проведения стереосъёмки спутники либо оснащаются двумя телескопами [1], либо выполняют конвергентную съёмку местности одним телескопом за счёт разворота космического аппарата (КА) по углу тангажа (КА типа «Ресурс-П», PLEIADES и др.). Угол, под которым выполняется наблюдение заданной территории с двух точек, называется углом конвергенции, значения которого задают в пределах от 40° до 60°.

Методы обработки стереоизображений, особенно от кадровых систем дистанционного зондирования, достаточно подробно рассмотрены в литературе [2, 3]. Их особенность состоит в том, что на основе известных координат опорных точек местности, идентифицированных на стереоснимках, выполняется уточнение элементов внешнего ориентирования камеры (угловой ориентации и пространственного положения для двух точек положения КА). После этого с использованием метода прямой фотограмметрической засечки определяются трёхмерные (3D) координаты объектов земной поверхности. Достижимые при этом точности определения плановых и высотных координат объектов, наблюдаемых на стереоснимках, соответствуют точности задания координат опорных точек, что позволяет решать задачу уточнения карт масштаба 1:25000 и крупнее.

*Рассматриваются проблемные вопросы, связанные с крупномасштабным картографированием территорий по данным космической стереосъёмки без использования опорных точек местности. Показано, что точность определения по снимкам высоты должна быть в несколько раз выше, чем плановых координат объектов местности. Проведённый анализ показывает, что из-за погрешностей определения элементов внешнего ориентирования спутника в расчетах высоты объектов присутствуют систематические ошибки. Приводятся результаты эксперимента, подтверждающие наличие систематических ошибок при формировании карты высот местности. Описывается алгоритм компенсации систематических ошибок с помощью измерений лазерного дальномера. Приводится формула метода, позволяющего получать уточнённую карту высот и на её основе ортоплана картографируемой территории, не обеспеченной опорными точками и цифровой моделью рельефа.*

К сожалению, для случаев, когда отсутствуют опорные точки для наблюдаемых территориях, известные решения оказываются не работоспособными. Поэтому актуальной является задача высокоточного определения трёхмерных координат объектов земной поверхности при отсутствии опорных точек местности. Для её решения проанализируем источники ошибок, возникающих при стереообработке сканерных снимков, и рассмотрим метод, позволяющий повысить точность определения высотных координат объектов и формировать высокоточные 3D-карты для любых территорий земной поверхности.

#### Модель обработки стереоизображений

Обработка двух изображений земной поверхности  $I_1(m_1, n_1)$  и  $I_2(m_2, n_2)$ ,  $m_1, m_2 \in [1, M]$ ,  $n_1, n_2 \in [1, N]$ , получаемых при стереосъёмке, выполняется на основе строгой модели функционирования съёмочной системы. Модель устанавливает для каждого снимка аналитическую зависимость между пиксельными  $(m_i, n_i)$ ,  $i = 1, 2$  и геодезическими  $(\varphi, \lambda)$  координатами в зависимости от

высоты местности  $h$ , линейных  $\mathbf{p}_i$ , угловых  $\mathbf{\Omega}_i$  элементов внешнего и внутреннего  $\mathbf{r}_i$  ориентирования, а также конструктивных углов камеры  $\mathbf{R}_i$ . Для случая однопленочной съёмки значения  $\mathbf{r}_1 = \mathbf{r}_2$  и  $\mathbf{R}_1 = \mathbf{R}_2$ . Процесс построения модели рассмотрен в [4], поэтому запишем искомую зависимость в виде обобщённых уравнений для случаев прямого и обратного координатного соответствия:

$$\begin{aligned} \varphi &= \Phi(m_i, n_i, h, \mathbf{R}_i, \mathbf{p}_i, \mathbf{\Omega}_i, \mathbf{r}_i), \\ \lambda &= F(m_i, n_i, h, \mathbf{R}_i, \mathbf{p}_i, \mathbf{\Omega}_i, \mathbf{r}_i), \\ m_i &= \Phi^{-1}(\varphi, \lambda, h, \mathbf{R}_i, \mathbf{p}_i, \mathbf{\Omega}_i, \mathbf{r}_i), \\ n_i &= F^{-1}(\varphi, \lambda, h, \mathbf{R}_i, \mathbf{p}_i, \mathbf{\Omega}_i, \mathbf{r}_i), \quad i = 1, 2. \end{aligned} \quad (1)$$

В рассматриваемой модели элементы внешнего ориентирования  $\mathbf{p}_i$ ,  $\mathbf{\Omega}_i$ ,  $i = 1, 2$ , задающие пространственное и угловое положение камеры, определяются по данным измерительных систем спутника с СКО  $\sigma_p$  и  $\sigma_{\Omega}$ . Это означает, что во время съёмки реальное положение визирного луча камеры отличается от идеального. При заданной высоте местности  $h$  ошибки  $\sigma_p$  и  $\sigma_{\Omega}$  приводят к ошибочному определению плановых координат объектов земной поверхности в геодезической  $\sigma_{\varphi}$  и  $\sigma_{\lambda}$  или картографической (метрической)  $\sigma_{II}$  системе координат:

$$\sigma_{II} = \left\{ [R_s \sigma_{\varphi}]^2 + [R_s \sigma_{\lambda} \cos \varphi]^2 \right\}^{1/2}, \quad (2)$$

где  $R_s$  – радиус земного эллипсоида,  $R_s = 6371$  км;  $\varphi$  – геодезическая широта точки снимка.

Исследования показывают, что для современных КА высокого пространственного разрешения («Ресурс-П», IKONOS, GeoEye1, WorldView) ошибки определения плановых координат  $\sigma_{II}$  находятся в пределах 6-12 м, что вполне соответствует требованиям по точности для карт масштаба 1:25000.

При стереобработке трёхмерные координаты объектов земной поверхности  $(\varphi, \lambda, h)$ , идентифицированных на снимках  $I_1$  и  $I_2$ , вычисляются путём решения переопределённой системы 4-х уравнений с тремя неизвестными:

$$\begin{cases} m_1 = \Phi^{-1}(\varphi, \lambda, h, \mathbf{R}_1, \mathbf{p}_1, \mathbf{\Omega}_1, \mathbf{r}_1), \\ n_1 = F^{-1}(\varphi, \lambda, h, \mathbf{R}_1, \mathbf{p}_1, \mathbf{\Omega}_1, \mathbf{r}_1), \\ m_2 = \Phi^{-1}(\varphi, \lambda, h, \mathbf{R}_2, \mathbf{p}_2, \mathbf{\Omega}_2, \mathbf{r}_2), \\ n_2 = F^{-1}(\varphi, \lambda, h, \mathbf{R}_2, \mathbf{p}_2, \mathbf{\Omega}_2, \mathbf{r}_2). \end{cases} \quad (3)$$

Нетрудно заметить, что СКО ошибки определения плановых координат объектов будет соответствовать значению, вычисленному по формуле (2). Что касается ошибки определения высоты объектов  $\sigma_h$ , то она дополнительно зависит от угла конвергенции  $\alpha$  и связана со значением  $\sigma_{II}$  известным соотношением [5],

$$\sigma_h = \frac{\sigma_{II}}{B/H}, \quad (4)$$

где  $B$  – величина стереобазы, определяемая углом конвергенции,  $H$  – высота съёмки.

Рис. 1 поясняет физический смысл формулы (4). Идеальное положение визирных лучей камеры представлено пунктирными линиями, а реальное положение визирного луча камеры из точки 2 показано непрерывной линией. Из-за различия углового положения лучей, координаты вершины здания определяются с ошибками  $\sigma_{II}$  и  $\sigma_h$ . При соотношении  $B/H \approx 1$  (угол конвергенции  $\alpha \approx 60^\circ$ ) значение  $\sigma_h = \sigma_{II}$  и составит величину порядка 6-12 м, что в несколько раз превосходит допустимую ошибку определения высот для карт масштаба 1:25000.

Из сказанного следует важный вывод. Ошибка определения плановых координат объектов  $\sigma_{II}$  слабо изменяется по полю снимка и, как правило, в пределах условного кадра является постоянной величиной (за условный кадр принимается изображение, в котором число строк и столбцов совпадают). Тогда, пренебрегая ошибкой идентификации одноимённых объектов на стереоизображениях, можно сделать заключение о наличии систематической ошибки определения высотных координат объектов земной поверхности.

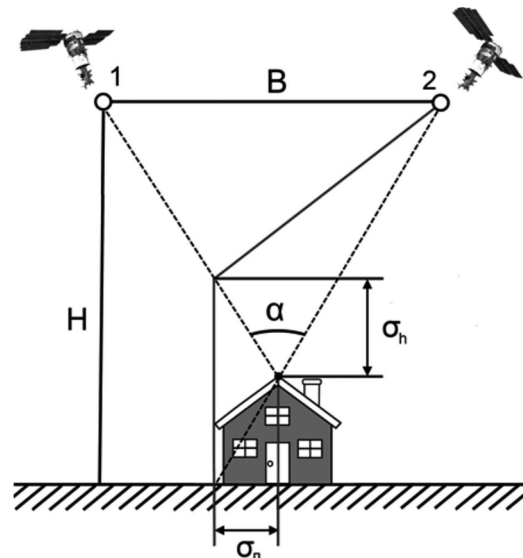
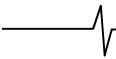


Рис. 1. Схема стереосъёмки земной поверхности

#### Эксперимент по подтверждению наличия систематических ошибок при определении высот объектов

Для эксперимента КА «Ресурс-П» № 1 была выполнена конвергентная стереосъёмка ( $\alpha = 50^\circ$ ) территории г. Самара, обеспеченной 40 опорными точками с известными геодезическими координатами и высотой  $\varphi_0(i)$ ,  $\lambda_0(i)$ ,  $h_0(i)$ ,  $i = \overline{1, 40}$  (маршруты 0041\_0102\_05164\_1\_05163\_03\_10.tif и 0041\_0102\_05164\_1\_05163\_02\_10.tif). С помощью специализированного программного комплекса ССП-ВР была выполнена стереобработка полученных маршрутов съёмки с использованием только измерительной информации. На основе соотношений (3) были вычислены трёхмерные координаты  $\varphi(i)$ ,  $\lambda(i)$ ,  $h(i)$  объектов, соответствующих изображениям опорных точек, и рассчитаны разности высот:  $\Delta h(i) = h(i) - h_0(i)$ ,  $i = \overline{1, 40}$ . Для дальнейшей статистической обработки результатов рассматривалась аддитивная модель оши-



бок, согласно которой разности высот в опорных точках рассматривались в виде суммы систематической  $\delta_h$  и случайной  $\varepsilon_h$  ошибок:

$$\Delta h = \delta_h + \varepsilon_h. \quad (5)$$

Среднее значение разности высот, оценка систематической ошибки, вычислялась как:

$$\overline{\delta_h} = \frac{1}{40} \sum_{i=1}^{40} \Delta h(i), \quad (6)$$

а среднеквадратическая ошибка  $\sigma_h$ :

$$\sigma_h = \sqrt{\frac{1}{40} \sum_{i=1}^{40} \Delta h^2(i)}. \quad (7)$$

После исключения систематической ошибки из результатов измерений  $\delta_h = \Delta h(i) - \overline{\delta_h}$ , оценивались параметры случайной составляющей как стандартной среднеквадратической ошибки:

$$\varepsilon_h = \sqrt{\frac{1}{39} \sum_{i=1}^{40} \delta_h^2(i)}. \quad (8)$$

Результаты расчётов следующие:  $\overline{\delta_h} = 9,01$  м,  $\sigma_h = 9,13$  м,  $\varepsilon_h = 1,4$  м. Малое значение случайной ошибки  $\varepsilon_h$  и близкие значения  $\sigma_h$  и  $\overline{\delta_h}$  подтверждают вывод о наличии систематической ошибки в вычислении координат высот объектов. Данное заключение согласовывается с исследованиями точности ЦМР, созданных по стереопарам китайского спутника TH-1 [6].

#### Алгоритм компенсации систематических ошибок расчёта высоты объектов по данным лазерного дальномера

Устранить систематические ошибки в определении высот можно, если, зная точное значение высоты  $h_0$  в одной или нескольких точках стереопары, вычислить поправочный коэффициент  $\Delta h$ , который затем учесть в расчетах. Получить точное значение высоты можно с использованием лазерного дальномера, которым оснащается КА [1]. Как отмечается в публикациях, лазерный дальномер облучает на земной поверхности область  $\Theta$  диаметром порядка 60-70 м, содержащую осреднённую информацию о всех находящихся в ней высотных объектах. При этом расстояния между соседними отсчётами на несколько порядков превышают пространственное разрешение стереоснимков. Понятно, что подобные измерения весьма грубы и не могут непосредственно использоваться для вычисления поправочного коэффициента и формирования высокоточной карты высот наблюдаемой территории.

Для обеспечения возможности использования измерений лазерного дальномера (ЛД) по стереопаре сформируем карту высот местности  $h(\varphi_i, \lambda_j)$  с шагом  $(\Delta\varphi, \Delta\lambda)$ ,  $\varphi_i = \varphi_{i-1} + \Delta\varphi$ ,  $\lambda_j = \lambda_{j-1} + \Delta\lambda$ . Затем рассчитаем среднее значение высоты  $\overline{h_\Theta}$ , соответствующей области  $\Theta$ . В этом случае значение  $\overline{h_\Theta}$  будет соответствовать значению высоты  $h_L$ , которую измеряет ЛД. Отсюда поправочный коэффициент  $\Delta h = \overline{h_\Theta} - h_L$ . С его

учётом уточнённая карта высот  $h^*(\varphi_i, \lambda_j)$  будет сформирована по правилу  $h^*(\varphi_i, \lambda_j) = h(\varphi_i, \lambda_j) + \Delta h$ .

Предложенный алгоритм определения поправочного коэффициента основан на модели функционирования ЛД. Для её построения введём следующие обозначения: переменной  $S$  обозначим наклонную дальность, измененную ЛД;  $\mathbf{R}_L$ ,  $\mathbf{\Omega}_L$  – матрицами установочные углы и угловые элементы внешнего ориентирования, а  $\gamma$  – угловое поле зрения ЛД. Тогда, по аналогии с моделью функционирования камеры [4], запишем уравнение для вычисления гринвичских координат центральной точки области  $\Theta$ ,

$$\mathbf{p}_r(t) = \mathbf{p}(t) + S(t)\mathbf{R}_r(t)\mathbf{\Omega}_L\mathbf{R}_L\mathbf{r}_L, \quad (9)$$

где  $\mathbf{p}_r(t) = [X(t), Y(t), Z(t)]^T$  – вектор гринвичских координат точки на земной поверхности,  $t$  – время формирования измерения ЛД,  $\mathbf{R}_r(t)$  – матрица перехода от связанной системы координат КА к гринвичской,

$\mathbf{r}_L = [0, 0, 1]^T$ . С помощью известных соотношений по трёхмерным координатам точки  $\mathbf{p}_r(t)$  рассчитаем геодезические координаты центра области  $\Theta$ , а затем, зная угловое поле зрения ЛД, нетрудно вычислить граничные значения  $[\varphi_{\min}(t), \lambda_{\min}(t)]$ ,  $[\varphi_{\max}(t), \lambda_{\max}(t)]$  области  $\Theta$ . Значение  $\overline{h_\Theta}$ , соответствующее текущему измерению ЛД, вычислим на основе карт высот  $h(\varphi_i, \lambda_j)$ ,

$$\overline{h_\Theta} = \frac{1}{I_\varphi J_\lambda} \sum_{i=1}^{I_\varphi} \sum_{j=1}^{J_\lambda} h(\varphi_i, \lambda_j), \quad (10)$$

$$I_\varphi = (\varphi_{\max}(t) - \varphi_{\min}(t)) / \Delta\varphi, J_\lambda = (\lambda_{\max}(t) - \lambda_{\min}(t)) / \Delta\lambda.$$

На основе анализа уравнения (9) получим априорную оценку измеренной ЛД высоты  $h_L$ . Так, при высокоточном (доли метра) измерении наклонной дальности  $S$  высота местности будет определяться с погрешностью измерения высоты КА. Как показали исследования [4] СКО измерения высоты КА приёмником ГЛОНАСС/GPS после наземной обработки составляет 2-3 м. Соответственно, с такой же точностью будет определяться поправочный коэффициент  $\Delta h$ . Таким образом, если бы при стереосъёмке Самарского полигона использовались измерения ЛД, то СКО определения высоты в опорных точках приняло значение  $\sigma_h \approx 3,4$  м, что вполне удовлетворило бы требованиям масштаба 1:25000.

#### Формула метода высокоточного картографирования по данным стереосъёмки и измерениям ЛД

Для высокоточного картографирования земной поверхности КА помимо съёмочной аппаратуры, представленной одной или двумя камерами, должен оснащаться ЛД, формирующим измерения в поле зрения камеры наблюдения. С помощью съёмочной аппаратуры выполняется стереосъёмка требуемой территории, в результате которой формируются два стереоизображения  $I_1(m_1, n_1)$ ,  $I_2(m_2, n_2)$  и измерения лазерного дальномера  $S(t)$ , рис. 2. С использованием уравнений (3) формирует-

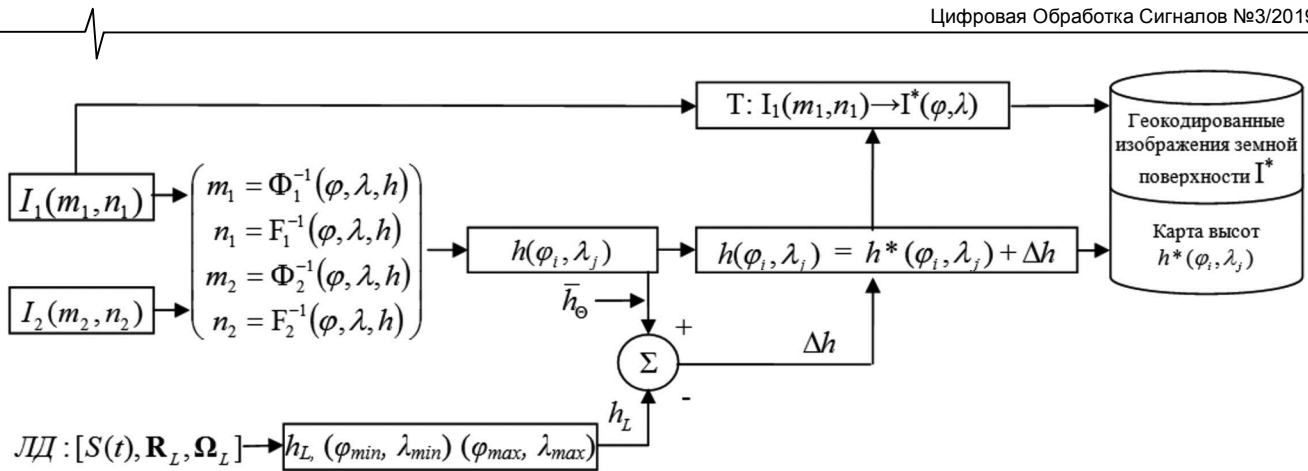


Рис. 2. Схема метода высокоточного картографирования

ся карта высот местности  $h(\varphi_i, \lambda_j)$ , а на основе формулы (9) вычисляются геодезические координаты области облучения земной поверхности дальномером  $\Theta$ . По карте высот  $h(\varphi_i, \lambda_j)$ , соответствующей области  $\Theta$ , вычисляется поправочный коэффициент  $\Delta h$ , компенсирующий систематические ошибки в расчётах высоты, на основе которого формируется уточнённый массив высот  $h^*(\varphi_i, \lambda_j)$  картографируемой территории. С учётом уточнённой цифровой модели местности  $h^*(\varphi_i, \lambda_j)$  и обратных уравнений координатного соответствия (1) выполняется ортотрансформирование одного из снимков стереопары в геодезическую или картографическую систему координат. Тем самым формируется геокодированное изображение земной поверхности, для которого минимизированы ошибки определения плановых координат объектов за счёт использования высокоточной цифровой модели местности  $h^*(\varphi_i, \lambda_j)$ .

### Заключение

На основе рассмотренного метода будут выполнены доработки программного обеспечения стереообработки изображений высокого разрешения от планируемого к запуску КА «Аист-2Т» и других. Поскольку в отечественной практике спутников, оснащённых ЛД нет, то отработка метода будет выполняться на натурных данных от перспективных КА. При этом многое будет зависеть от точности работы ЛД и его установки на спутнике. В слу-

чае положительных результатов натурных испытаний для топографов будет предоставлена информация для создания востребованных на практике карт масштаба 1:25000.

### Литература

1. Кириллин А.Н., Ахметов Р.Н., Салмин В.В. и др. Формирование на базе платформы «Аист-2» проектного облика малого космического аппарата обзорного наблюдения, функционирующего в составе орбитальной группировки: С. 32-35 / Системы наблюдения, мониторинга и дистанционного зондирования Земли. Материалы XV научно-технической конференции (респ. Крым, г. Алушта, 16-21 сентября 2018 г.). Калуга: Манускрипт, 2018. – 368 с.
2. Лобанов А.Н. Фотограмметрия – М., Недра, 1984, 552 с.
3. Назаров А.С. Фотограмметрия: пособия для студентов вузов – Минск: ТерраСистемс, 2010, 400 с.
4. Ахметов Р.Н., Еремеев В.В., Кузнецов А.Е., Мятов Г.Н., Пошехонов В.И., Стратилатов Н.Р. Высокоточная геодезическая привязка изображений земной поверхности от КА «Ресурс-П» // Исследование Земли из космоса, №1, 2017, с. 44-53.
5. Dowman I., Jacobsen K., Konecny G., Sandau R. High Resolution Optical Satellite Imagery. Whittles Publishing, 2011. 230 p.
6. Оньков И.В. Оценка точности ЦМР, созданных по стереопарам триплета КА ТН-1 в программном комплексе ENVI // Геоматика. – 2014. – № 3. с. 22-27.

## МЕТОДИКА ФОРМИРОВАНИЯ КОЭФФИЦИЕНТОВ ВАЖНОСТИ ОБЪЕКТОВ, НАБЛЮДАЕМЫХ НА ИЗОБРАЖЕНИЯХ СРЕДСТВАМИ ДИСТАНЦИОННОГО ЗОНДИРОВАНИЯ ЗЕМЛИ

*Мятов Г.Н., д.т.н., заместитель главного конструктора АО «РКЦ «Прогресс», e-mail: miatov@mail.ru;  
Поливанов В.А., к.т.н., ВА РВСН.*

### CREATION TECHNIQUE OF CRUCIAL OBJECT COEFFICIENT OBSERVED IN SATELLITE IMAGES

*Miatov G.N., Polivanov V.A.*

*In this paper technique of effective satellite constellation imagery of the area is presented. The problem is solved by selection crucial object coefficient allowed distribute resource of satellite survey in constellation optimally.*

**Key words:** satellite constellation, Earth remote sensing, geodetic reference of images, plan of satellite survey.

**Ключевые слова:** орбитальная группировка, дистанционное зондирование земли, плановая ситуация съемки, объект наблюдения (съемки).

#### Введение

Увеличение спектра задач, решаемых средствами дистанционного зондирования Земли (ДЗЗ), количества потребителей информации ДЗЗ, а также номенклатуры КА ДЗЗ делает проблему планирования целевого применения орбитальных группировок (ОГ) ДЗЗ особенно актуальной. А учитывая тенденцию кратного роста в ближайшие годы количества КА ДЗЗ на орбите – критической.

Ключевым фактором, обеспечивающим эффективное решение задач ДЗЗ имеющейся ОГ, является подбор таких коэффициентов важности объектов (районов) наблюдения, которые позволят распределить ресурс средств ДЗЗ наиболее эффективно. При этом накладывается условие, которое учитывает плановую ситуацию [1, 2].

#### Постановка задачи

Под плановой ситуацией будем понимать перечень задач, которые надо решить, и объектов, съемку которых необходимо осуществить для решения данных задач, с указанием:

- требуемого срока и качества выполнения задач (съемки объектов);
- требуемого количества наблюдений каждого объекта за указанный интервал времени;
- относительной важности потребителя каждой задачи (объекта).

Рассмотрим алгоритм формирования коэффициентов важности объектов (районов) наблюдения, входящих в задачи, решаемые средствами ДЗЗ.

Определим степень (уровень) важности каждой задачи как совокупность:

- уровня лица (потребителя), поставившего задачу;

*Рассматривается методика формирования коэффициентов важности объектов, обеспечивающая эффективное решение задач съемки интересующих районов орбитальной группировкой (ОГ) дистанционного зондирования Земли (ДЗЗ), путем подбора таких коэффициентов важности объектов (районов) наблюдения, которые позволяют распределить ресурс средств ДЗЗ наиболее эффективно.*

- требуемого конечного срока выполнения задачи.

Обозначим  $Z(Z_1, Z_2, \dots, Z_N)$  – вектор задач на заданный период времени,  $O(O_1, O_2, \dots, O_M)$  – вектор объектов, съемку которых необходимо осуществить с заданными требованиями для решения перечня задач.

Каталог задач представим графом с вершинами в виде задач и объектов (районов), а также связями между ними (рис. 1).

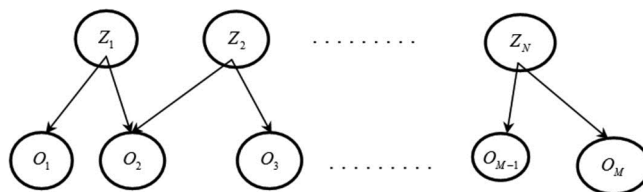


Рис. 1. Представление каталога задач в виде графа

Тогда графическую иллюстрацию изменения важности каждого объекта можно представить в виде графиков, приведенных на рис. 2.

На рис. 2 приняты следующие обозначения:

- $b_1, b_2, b_N$  – уровень лиц, поставивших задачи  $Z_1, Z_2, Z_N$  соответственно;
- $1, 2, N$  – требуемое количество съемок объектов задач  $Z_1, Z_2, Z_N$  соответственно за интервал времени  $t_k - t_0$ ;
- $Y_{\max}$  – максимально допустимое значение коэффициента важности для всех объектов всех задач.

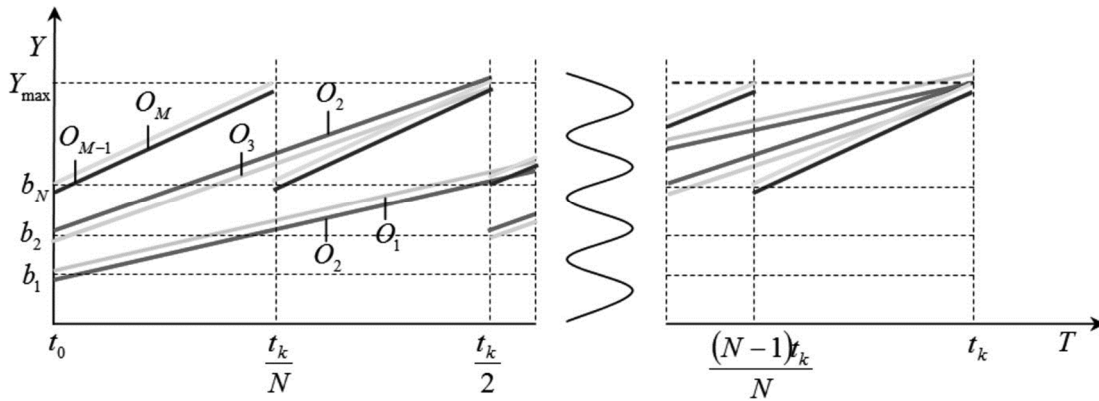


Рис. 2. Графическая иллюстрация изменения важности объектов

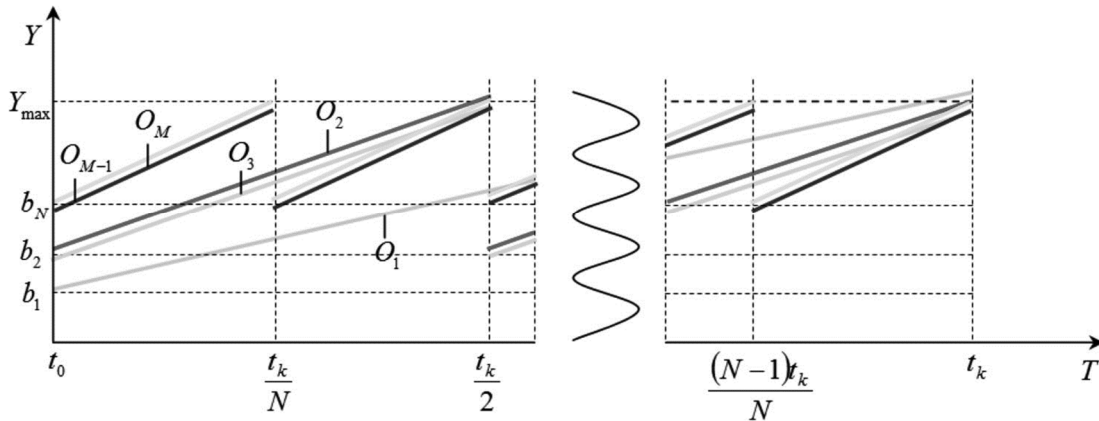


Рис. 3. Графическая иллюстрация изменения важности объектов плановой ситуации уменьшенной размерности

Множество функций, определяющих изменение коэффициентов важности всех объектов всех задач, определим следующим образом:

$$\begin{cases}
 y_1 = a_1x_1 + b_1 - k(Y_{\max} - b_1), \\
 \text{для объекта } O_1 \text{ задачи } Z_1 \\
 y_2 = a_1x_2 + b_1 - k(Y_{\max} - b_1), \\
 \text{для объекта } O_2 \text{ задачи } Z_1 \\
 y_3 = a_2x_3 + b_2 - k(Y_{\max} - b_2), \\
 \text{для объекта } O_2 \text{ задачи } Z_2 \\
 y_4 = a_2x_4 + b_2 - k(Y_{\max} - b_2), \\
 \text{для объекта } O_3 \text{ задачи } Z_2 \\
 \text{-----} \\
 y_{M+S} = a_Nx_{M+S} + b_N - k(Y_{\max} - b_N), \\
 \text{для объекта } O_M \text{ задачи } Z_N,
 \end{cases} \quad (1)$$

$$k = 0, 1, \dots, (N - 1),$$

где  $k$  – коэффициент, определяющий номер интервала наблюдения.

С целью оптимизации решения задачи реализации плановой ситуации, снизим размерность системы функций, определяющих изменение коэффициентов важности всех объектов всех задач. Для этого из плановой ситуации удалим объекты, входящие одновременно в несколько задач (объект  $O_2$ ), из задачи, где требуется наименьшая периодичность съемки. Тогда система функций, определяющих изменение коэффициентов важности всех объектов всех задач, будет выглядеть следующим образом:

$$\begin{cases}
 y_1 = a_1x_1 + b_1 - k(Y_{\max} - b_1), \\
 \text{для объекта } O_1 \text{ задачи } Z_1 \\
 y_3 = a_2x_3 + b_2 - k(Y_{\max} - b_2), \\
 \text{для объекта } O_2 \text{ задачи } Z_2 \\
 y_4 = a_2x_4 + b_2 - k(Y_{\max} - b_2), \\
 \text{для объекта } O_3 \text{ задачи } Z_2 \\
 \text{-----} \\
 y_{M+S-S^1} = a_Nx_{M+S-S^1} + b_N - k(Y_{\max} - b_N), \\
 \text{для объекта } O_M \text{ задачи } Z_N,
 \end{cases} \quad (2)$$

$$k = 0, 1, \dots, (N - 1),$$

где  $k$  – коэффициент, определяющий номер интервала наблюдения.

Графическая иллюстрация изменения важности объектов представлена на рис. 3:

Рассмотрим функцию изменения коэффициента важности  $k$ -го объекта. Примем, что функция изменения коэффициента важности линейная вида:

$$y(t) = at + b,$$

где:

- $y$  – текущий коэффициент важности объекта (района);
- $a$  – коэффициент, зависящий от требуемого срока и периодичности наблюдения объекта (района);
- $t$  – текущий момент времени;
- $b$  – уровень лица, поставившего задачу, т.е.

$$b = UL.$$

Определим значение коэффициента  $a$ . Пусть:

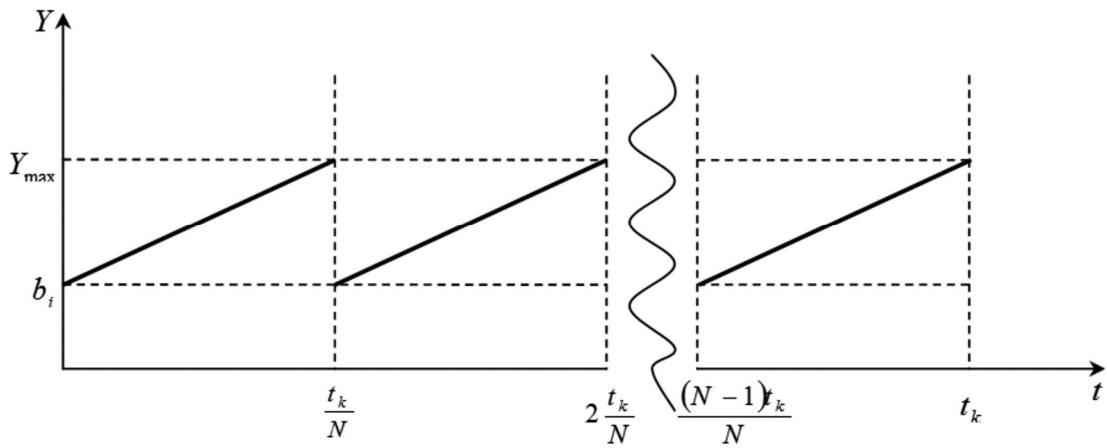


Рис. 4. Графическая иллюстрация изменения коэффициента важности

- $b_i$  – уровень  $i$ -го лица (потребителя), поставившего задачу;
- $t_0$  – начальный момент времени;
- $t_k$  – требуемый срок съемки  $k$ -го объекта;
- $N$  – требуемое количество съемок  $k$ -го объекта за интервал времени  $t_k - t_0$ ;
- $Y_{\max}$  – максимально допустимое значение коэффициента важности  $k$ -го объекта.

Изменение коэффициента важности  $y_k$   $k$ -го объекта можно отобразить графиком, представленным на рис. 4.

Тогда для любого момента времени  $t$  значение коэффициента  $a$  будет вычисляться по формуле:

$$a = \frac{(Y_{\max} - b_i)}{(t_k - t_0)} N.$$

Тогда

$$y(t) = \frac{(Y_{\max} - b_i)N}{(t_k - t_0)} t + b_i.$$

Рассмотрим различные варианты плановой ситуации, сформированной из задач и объектов данного графа.

Пусть каталог задач представлен следующим графом (рис. 5).

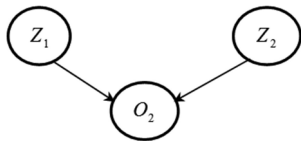


Рис. 5. Представление каталога задач в виде графа

При этом для задачи  $Z_1$  величина  $N = 1$ , а для задачи  $Z_2$  величина  $N = 2$ .

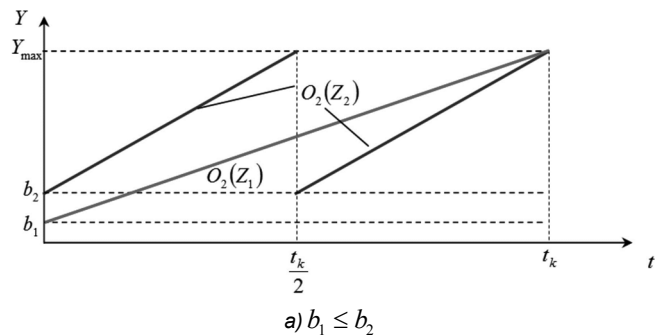
Пусть изменение коэффициента важности объекта  $O_2$  для задачи  $Z_1$  будет описываться функцией:

$$y(t) = a_1 t + b_1,$$

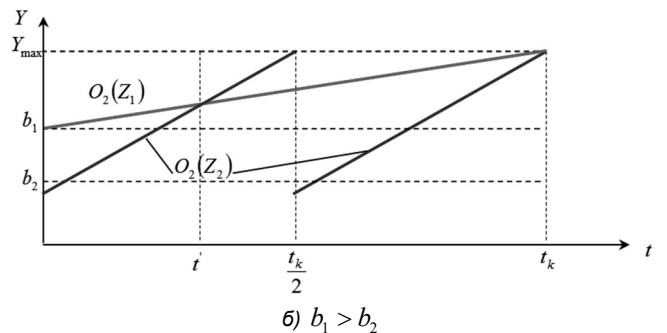
а объекта  $O_2$  для задачи  $Z_2$  функцией:

$$y(t) = a_2 t + b_2.$$

Возможны следующие варианты плановой ситуации (рис. 6 а и 6 б).



а)  $b_1 \leq b_2$



б)  $b_1 > b_2$

Рис. 6. Графическая иллюстрация изменения важности объектов

Наибольший интерес представляет плановая ситуация для случая, где  $b_1 > b_2$ . После снижения размерности системы функций, определяющих изменение коэффициентов важности объекта  $O_2$  задач  $Z_1$  и  $Z_2$ , в соответствии с алгоритмом снижения размерности, приведенным выше, графическая иллюстрация изменения важности объектов примет вид, представленный на рис. 7.

Функция, определяющая изменение коэффициента важности объекта  $O_2$  задач  $Z_1$  и  $Z_2$  на интервале

$\left[ t_0, \frac{t_k}{2} \right]$  будет выглядеть следующим образом:

$$\begin{cases} y(t) = \frac{(Y_{\max} - b_1)}{(t_k - t_0)} t + b_1, t \in \left[ t_0, \frac{(b_2 - b_1)(t_k - t_0)}{2b_2 - b_1 - Y_{\max}} \right] \\ y(t) = \frac{2(Y_{\max} - b_2)}{(t_k - t_0)} t + b_2, t \in \left( \frac{(b_2 - b_1)(t_k - t_0)}{2b_2 - b_1 - Y_{\max}}, t_k \right] \end{cases} \quad (3)$$



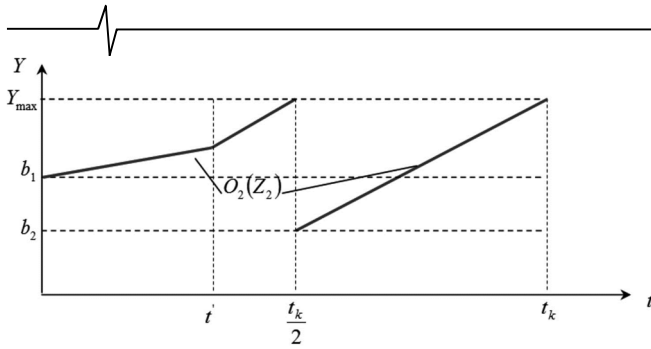


Рис. 7. Графическая иллюстрация изменения важности объектов плановой ситуации уменьшенной размерности для случая, когда  $b_1 > b_2$

Таким образом, реализация методики формирования коэффициентов важности объектов (районов) наблюдения сводится к выполнению последовательности действий:

1. Формирование на заданный период времени вектора задач  $Z(Z_1, Z_2, \dots, Z_N)$ .
2. Формирование на заданный период времени вектора объектов, съемку которых необходимо осуществить с заданными требованиями для решения перечня задач  $O(O_1, O_2, \dots, O_M)$ .
3. Формирование вектора  $B(b_1, \dots, b_N)$  – уровня лиц, поставивших задачи  $Z(Z_1, Z_2, \dots, Z_N)$ .
4. Формирование системы функций, определяющих изменение коэффициентов важности всех объектов всех задач, согласно уравнению (1).
5. Анализ плановой ситуации на всем интервале пла-

нирования  $[t_0, t_k]$  и уменьшение размерности системы функций, определяющих изменение коэффициентов важности всех объектов всех задач, согласно уравнению (2).

6. Анализ плановой ситуации на интервале планирования  $\left[ t_0, \frac{t_k}{N} \right]$  и формирование на этом интервале си-

стемы функций, определяющих изменение коэффициентов важности всех объектов всех задач согласно уравнению (3).

### Заключение

Таким образом, разработана методика, обеспечивающая эффективное решение задач съемки интересных районов имеющейся ОГ ДЗЗ, путем подбора таких коэффициентов важности объектов (районов) наблюдения, которые позволяют распределить ресурс средств ДЗЗ наиболее эффективно.

### Литература

1. Безняков А.М., Горбулин В.И., Каргу Д.Л., Фадеев А.С., Щербаков В.И. Методика оптимального планирования применения космических аппаратов Дистанционного Зондирования Земли в обзорном режиме // Вестник Самарского государственного аэрокосмического университета. 2013. № 1(39). С. 39–45.
2. Бузуев К.В. Определение оптимальных планов наблюдения космического аппарата дистанционного зондирования Земли с помощью графа // Вестник СГАУ им. академика С.П. Королёва. 2013. Вып. 1(39). С. 63-72.

## ВЛИЯНИЕ ТОЧНОСТИ ОПРЕДЕЛЕНИЯ УГЛОВЫХ ЭЛЕМЕНТОВ ВНЕШНЕГО ОРИЕНТИРОВАНИЯ НА ТОЧНОСТЬ ОЦЕНКИ КООРДИНАТ ОБЪЕКТОВ НА КОСМИЧЕСКИХ ИЗОБРАЖЕНИЯХ

*Бузев К.В., начальник сектора АО «РКЦ «Прогресс»;*

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

*Платошин И.В., ведущий инженер АО «РКЦ «Прогресс».*

## THE INFLUENCE OF ANGULAR EXTERNAL ORIENTATION ELEMENTS PRECISION ON GEODETIC REFERENCE OF IMAGES ACCURACY

*Buzuev K.V., Miatov G.N., Platoshin I.V.*

*The problem of estimating the influence of angular external elements on geolocation of satellite images is discussed. The problem solution via search algorithm and mathematical simulation is presented.*

**Key words:** satellite, remote sensing images, geodetic reference of images, angular external orientation elements.

**Ключевые слова:** космический аппарат, дистанционное зондирование Земли, точность оценки координат объекта, точность определения угловых элементов внешнего ориентирования, алгоритм, обратная геодезическая задача.

### Введение

В настоящее время космические аппараты (КА), предназначенные для дистанционного зондирования Земли (ДЗЗ) в видимом и ближнем инфракрасном диапазоне, например, такие как КА «Ресурс-П», применяются для решения следующих научно-технических задач [1]:

- составление и обновление общегеографических, тематических и топографических карт;
- контроль загрязнения и деградации окружающей среды, в том числе, экологический контроль в районах геологоразведочных работ и добычи полезных ископаемых, выявление и изучение загрязнений окружающей среды, контроль водоохранных и заповедных районов;
- наблюдение районов чрезвычайных ситуаций с целью предупреждения развития стихийных бедствий, аварий, катастроф, а также оценка их последствий с целью планирования восстановительных мероприятий;
- контроль застройки территорий, получение данных для инженерной оценки местности в интересах хозяйственной деятельности и многих других.

Каждая из этих задач накладывает определенные требования к основным тактико-техническим характеристикам КА ДЗЗ: линейному разрешению на местности, производительности КА, оперативности доставки информации потребителю и точности определения координат объектов на снимке.

Для решения задач создания и обновления топографических карт к КА «Ресурс-П» впервые были предъявлены требования к обеспечению привязки снимков с точностью СКО 10-15 м.

Один из параметров, который оказывает существенное влияние на точность определения координат объек-

*Рассматривается задача оценки влияния точности определения угловых элементов внешнего ориентирования на точность оценки координат объекта, находящегося на пересечении центральной линии визирования с общеземным эллипсоидом. Рассматривается решение задачи с помощью поискового алгоритма. Определены условия математического моделирования. В соответствии с предложенным алгоритмом проведена оценка влияния точности определения угловых элементов внешнего ориентирования на точность определения координат объекта для определенных условий.*

та (ТОКО) на снимке является точность оценки угловых элементов внешнего ориентирования (УЭВО) КА [2].

### Постановка задачи

В статье решается задача оценки влияния точности определения УЭВО на ТОКО, находящегося на пересечении центральной линии визирования с общеземным эллипсоидом.

Для ее решения предлагается следующий алгоритм:

- 1) Из определенного диапазона околокруговых солнечно-синхронных орбит, как основного класса орбит для КА оптико-электронного наблюдения (ОЭН), выбирается орбита с заданной средней высотой.
- 2) Для выбранной орбиты проводится интегрирование уравнений движения центра масс КА на интервале одного витка его полета с фиксированным шагом интегрирования.
- 3) На каждом шаге интегрирования (соответствующем моменту времени  $t_i \in [0; T_{\text{оск}}]$ , где  $T_{\text{оск}}$  – оскулирующий период обращения КА по орбите) выполняется построение векторов дальности ( $\vec{D}_j$ ) до поверхности общеземного эллипсоида (ОЗЭ) по поверхности конуса наведения (рис. 1), определяемого углом отклонения линии визирования от надира ( $\xi$ ) и углом отклонения линии визирования от курса ( $\zeta$ ). Здесь  $\xi$  – угол между радиус-вектором центра масс КА, взятым с обратным знаком, и вектором дальности до наблюдаемой точки,  $\zeta$  – угол между проекцией вектора дальности на плос-

кость  $OX_0Z_0$  и направлением оси  $OX_0$  орбитальной системы координат (ОСК). За положительное направление  $\zeta$  принимается вращение по часовой стрелке, если смотреть с положительного направления оси  $OY_0$ . Для каждого вектора  $\vec{D}_j$  находятся координаты точки  $M_{j0}^i$  пересечения вектора дальности и поверхности ОЗЭ –  $(\varphi_{j0}^i; \lambda_{j0}^i)$ .

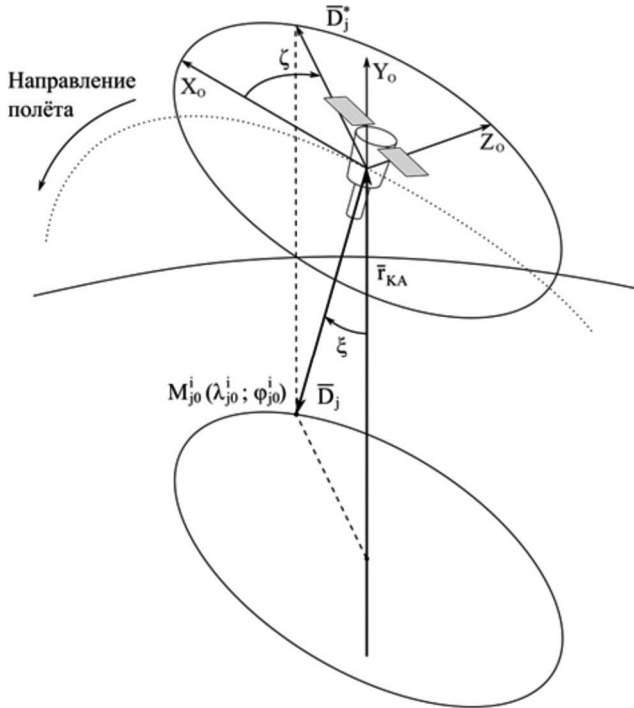


Рис. 1. Конус наведения КА

Для каждого вектора дальности строится набор векторов вспомогательных линий ( $\vec{A}_k$ ) по поверхности конуса рассеяния (рис. 2), определяемого углом отклонения вспомогательной линии от линии визирования ( $\Delta\xi$ ) и углом отклонения вспомогательной линии от курса ( $\Delta\zeta$ ). Здесь  $\Delta\xi$  – угол между вектором дальности и вектором вспомогательной линии,  $\Delta\zeta$  – угол между проекцией вектора вспомогательной линии на плоскость, перпендикулярную вектору дальности и направлением оси  $OX_D$  вспомогательной системы координат (ВспСК). Центр ВспСК находится в центре масс КА, ось  $OY_D$  противоположно направлена вектору дальности, ось  $OX_D$  лежит в плоскости, образованной осями  $OY_D$  и  $OY_0$ , перпендикулярна оси  $OY_D$  и направлена в сторону оси  $OY_0$  (для случая  $\xi = 0$  принимается, что ось  $OX_D$  совпадает с осью  $OX_0$ ), ось  $OZ_D$  дополняет систему до правой ортогональной. Для каждого вектора  $\vec{A}_k$  находятся координаты точки  $M_{jk}^i$  пересечения вспомогательной линии и поверхности ОЗЭ –  $(\varphi_{jk}^i; \lambda_{jk}^i)$ .

5. Находится расстояние  $L_{jk}^i$  между точками  $M_{jk}^i$  и  $M_{j0}^i$ .

6. Находится точность определения координат объекта для вектора дальности  $\vec{D}_j$  на момент времени  $t_i$ :

$$L_j^i = \max\{L_{jk}^i\}.$$

7. Находится точность определения координат объекта на момент времени  $t_i$ :  $L^i = \max\{L_j^i\}$ .

8. Находится точность определения координат объекта для текущей рабочей орбиты:  $L_{H=X} = \max\{L^i\}$ .

Для реализации предложенного алгоритма построена математическая модель.

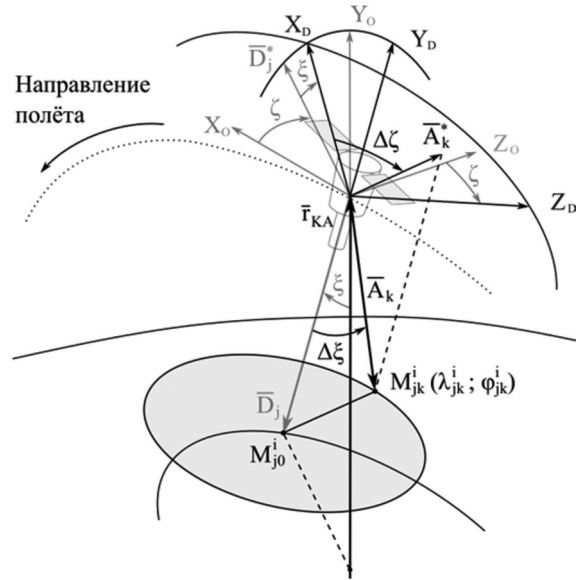


Рис. 2. Конус рассеяния

**Построение вектора дальности и вспомогательной линии**

Единичный вектор дальности может быть записан в проекции на оси ОСК в следующем виде:

$$\vec{l}_D^{OSK} = \frac{\vec{D}}{|\vec{D}|} = (\cos \zeta \cdot \sin \xi \quad -\cos \xi \quad \sin \zeta \cdot \sin \xi)^T. \tag{1}$$

Единичный вектор вспомогательной линии может быть записан в проекции на оси ВспСК в виде:

$$\vec{l}_A^{VspCK} = \frac{\vec{A}}{|\vec{A}|} = (\cos \Delta \zeta \cdot \sin \Delta \xi \quad -\cos \Delta \xi \quad \sin \Delta \zeta \cdot \sin \Delta \xi)^T. \tag{2}$$

Матрица перехода от ОСК к ИСК записывается в виде  $M_{OI} = (\vec{X}_0 | \vec{Y}_0 | \vec{Z}_0)$ , где  $X_0, Y_0, Z_0$  – оси ОСК, вычисляемые по формулам [3]:

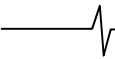
$$\vec{Y}_0 = \frac{\vec{r}}{|\vec{r}|}, \quad \vec{Z}_0 = -\frac{\vec{r} \times \vec{V}}{|\vec{r} \times \vec{V}|}, \quad \vec{X}_0 = \vec{Y}_0 \times \vec{Z}_0. \tag{3}$$

Матрица перехода от ВспСК к ОСК записывается в виде  $M_{DO} = (\vec{X}_D | \vec{Y}_D | \vec{Z}_D)$ , где  $X_D, Y_D, Z_D$  – оси ВспСК, вычисляемые по формулам:

$$\vec{Y}_D = -\frac{\vec{D}}{|\vec{D}|} = -\vec{l}_D^{OSK}, \quad \vec{Z}_D = -\frac{\vec{Y}_D \times \vec{Y}_0^{OSK}}{|\vec{Y}_D \times \vec{Y}_0^{OSK}|}, \tag{4}$$

$$\vec{X}_D = \frac{\vec{Y}_D \times \vec{Z}_D}{|\vec{Y}_D \times \vec{Z}_D|}, \quad \vec{Y}_0^{OSK} = (0 \quad 1 \quad 0)^T.$$

Проекции единичного вектора дальности на оси ИСК вычисляются как



$$\vec{l}_D^{ИСК} = M_{OH} \cdot \vec{l}_D^{ОСК} \quad (5)$$

Проекции единичного вектора вспомогательной линии на оси ИСК вычисляются по формуле:

$$\vec{l}_A^{ИСК} = M_{OH} \cdot M_{DO} \cdot \vec{l}_A^{БенСК} \quad (6)$$

### Нахождение координат точки пересечения линии визирования с ОЗЭ

Для нахождения радиус-вектора ( $\vec{R}_3$ ), исходящего из центра ОЗЭ в точку пересечения линии визирования с ОЗЭ, представим его как сумму радиус-вектора центра масс КА ( $\vec{r}_{КА}$ ) и вектора дальности  $\vec{D}$ :

$$\vec{R}_3 = \vec{r}_{КА} + \vec{D} \quad (7)$$

и подставим в уравнение ОЗЭ:

$$R_{3X}^2 + R_{3Y}^2 + \frac{R_{3Z}^2}{(1 - \alpha_{СЖ})^2} = R_3^2, \quad (8)$$

где  $R_{3X}$ ,  $R_{3Y}$ ,  $R_{3Z}$  – проекции радиус-вектора  $\vec{R}_3$  на оси ИСК,  $R_3 = 6378,136$  км – экваториальный радиус,

$\alpha_{СЖ} = \frac{1}{298,25784}$  – коэффициент полярного сжатия

ОЗЭ [4]. В результате получим уравнение:

$$(r_X + D_X)^2 + (r_Y + D_Y)^2 + \frac{(r_Z + D_Z)^2}{(1 - \alpha_{СЖ})^2} = R_3^2, \quad (9)$$

где  $r_X, r_Y, r_Z, D_X, D_Y, D_Z$  – проекции векторов  $\vec{r}_{КА}$  и  $\vec{D}$  на оси ИСК.

Раскроем скобки и проведем группировку слагаемых относительно модуля вектора дальности ( $D$ ), учитывая, что  $D_X = D \cdot l_{DX}^{ИСК}$ ,  $D_Y = D \cdot l_{DY}^{ИСК}$ ,  $D_Z = D \cdot l_{DZ}^{ИСК}$ . Получим уравнение:

$$\begin{aligned} & \left( l_{DX}^2 + l_{DY}^2 + \frac{l_{DZ}^2}{(1 - \alpha_{СЖ})^2} \right) \cdot D^2 + \\ & + 2 \cdot \left( r_X \cdot l_{DX}^{ИСК} + r_Y \cdot l_{DY}^{ИСК} + \frac{r_Z \cdot l_{DZ}^{ИСК}}{(1 - \alpha_{СЖ})^2} \right) \cdot D + \\ & + \left( r_X^2 + r_Y^2 + \frac{r_Z^2}{(1 - \alpha_{СЖ})^2} \right) - R_3^2 = 0, \end{aligned} \quad (10)$$

где

$$l_{DX}^2 = l_{DX}^{ИСК} \cdot l_{DX}^{ИСК}, \quad l_{DY}^2 = l_{DY}^{ИСК} \cdot l_{DY}^{ИСК}, \quad l_{DZ}^2 = l_{DZ}^{ИСК} \cdot l_{DZ}^{ИСК}, \quad l_{DX}^{ИСК}, l_{DY}^{ИСК}, l_{DZ}^{ИСК} - \text{проекции вектора } \vec{l}_D^{ИСК} \text{ на оси ИСК.}$$

Введем следующие обозначения:

$$\vec{r}_{II} = \left( r_X \quad r_Y \quad \frac{r_Z}{(1 - \alpha_{СЖ})} \right)^T - \text{приведенный радиус-вектор КА;}$$

вектор КА;

$$\vec{l}_{II} = \left( l_{DX}^{ИСК} \quad l_{DY}^{ИСК} \quad \frac{l_{DZ}^{ИСК}}{(1 - \alpha_{СЖ})} \right)^T - \text{приведенный единичный вектор дальности;}$$

вектор дальности;

$$C_{rI} = \vec{r}_{II} \cdot \vec{l}_{II} - \text{скалярное произведение векторов } \vec{r}_{II} \text{ и } \vec{l}_{II}.$$

Тогда уравнение (10) запишется в виде:

$$l_{II}^2 \cdot D^2 + 2 \cdot C_{rI} \cdot D + (r_{II}^2 - R_3^2) = 0. \quad (11)$$

Корни этого квадратного уравнения находятся в виде:

$$D_{1,2} = \frac{-C_{rI} \pm \sqrt{C_{rI}^2 + l_{II}^2 \cdot (R_3^2 - r_{II}^2)}}{l_{II}^2}. \quad (12)$$

Введем обозначения:  $d = C_{rI}^2 + l_{II}^2 \cdot (R_3^2 - r_{II}^2)$ ,

$$a = \frac{-C_{rI}}{l_{II}^2}, \quad b = \frac{\sqrt{d}}{l_{II}^2}. \text{ Отметим, что } a > 0 \text{ и } b > 0. \text{ Для слу-}$$

чая  $d > 0$  (пересечение линией визирования поверхности ОЗЭ) формула (12) запишется как  $D_1 = a - b$ ,  $D_2 = a + b$ , а искомым корнем будет  $D_1$ .

Тогда геоцентрические координаты точки пересечения линии визирования с поверхностью ОЗЭ находятся по формуле:

$$\begin{cases} \varphi = \arcsin \frac{R_{3Z}}{R_3} \\ \lambda = \arcsin \frac{R_{3X}}{R_3 \cdot \cos \varphi}, \text{ sign}(\cos \lambda) = \text{sign}(R_{3X}), \end{cases} \quad (13)$$

где  $R_{3i} = r_i + D_1 \cdot l_{Di}^{ИСК}$ ,  $i = X, Y, Z$ , а

$$R_3 = \sqrt{R_{3X}^2 + R_{3Y}^2 + R_{3Z}^2}.$$

Геоцентрические координаты точки пересечения вспомогательной линии с поверхностью ОЗЭ находятся аналогично, путем замены в формулах (10)-(13) вектора  $\vec{l}_D^{ИСК}$  на вектор  $\vec{l}_A^{ИСК}$ , вычисленный по формуле (6).

Перевод геоцентрической широты в геодезическую (долготы в геоцентрической и геодезической системах координат совпадают) осуществляется по формуле:

$$\text{tg} B = \frac{\text{tg} \varphi}{(1 - \alpha_{СЖ})^2}. \quad (14)$$

### Нахождение расстояния между точками пересечения $\vec{D}_j$ и $\vec{A}_k$ с поверхностью ОЗЭ

Нахождение расстояния между точками пересечения линии визирования и вспомогательной линии с поверхностью ОЗЭ является частью решения обратной геодезической задачи на малые расстояния по формулам со средними аргументами и проводится в соответствии с [5] по формулам:

$$b = \frac{B_2 - B_1}{\rho''}; \quad l = \frac{L_2 - L_1}{\rho''}, \quad (15)$$

где  $b$  – разность геодезических широт, выраженная в радианах,  $l$  – разность долгот, выраженная в радианах,

$\rho'' = \frac{180}{\pi} \cdot 3600$  – коэффициент перевода угловых секунд в радианы;

кунд в радианы;

$$B_m = \frac{B_1 + B_2}{2}; \quad \eta_m^2 = e'^2 \cdot \cos^2 B_m, \quad (16)$$

где  $B_m$  – средняя широта,  $\eta_m$  – вспомогательная функ-

ция эксцентриситета для средней точки,  $e' = \frac{\sqrt{R_3^2 - R_{II}^2}}{R_{II}}$  –

второй эксцентриситет,  $R_{II} = 6356,75136$  км – полярный радиус ОЗЭ;

$$Q = b \cdot M_m \cdot \left( 1 - (e'^2 - 2 \cdot \eta_m^2) \frac{R_{II}^2}{8} - (1 + \eta_m^2) \cdot \frac{(l \cdot \cos B_m)^2}{12} - \frac{(l \cdot \sin B_m)^2}{8} \right), \quad (17)$$

$$P = l \cdot \cos B_m \cdot N_m \times \left( 1 + (1 - 9 \cdot e'^2 + 8 \cdot \eta_m^2) \frac{b^2}{24} - \frac{(l \cdot \sin B_m)^2}{24} \right), \quad (18)$$

где  $M_m$  и  $N_m$  – главные радиусы кривизны. Эти параметры вычисляются в средней точке по формулам:

$$M_m = \frac{R_{\text{Э}} \cdot (1 - e^2)}{\sqrt{(1 - e^2 \cdot \sin^2 B_m)^3}}, \quad N_m = \frac{R_{\text{Э}}}{\sqrt{1 - e^2 \cdot \sin^2 B_m}},$$

$$e = \frac{\sqrt{R_{\text{Э}}^2 - R_{II}^2}}{R_{\text{Э}}}. \quad (19)$$

Искомое кратчайшее расстояние ( $s$  – длина геодезической линии) между точками находится как

$$s = \sqrt{Q^2 + P^2}. \quad (20)$$

### Исходные данные и результаты моделирования

В качестве исходных данных для проведения оценки были выбраны:

- около круговые солнечно-синхронные орбиты со средней высотой от 300 до 800 км с шагом 50 км;
- количество расчетных точек для каждой орбиты: 100 точек;
- распределение точек по орбите: равномерное;
- углы отклонения линии визирования от надира ( $\xi$ ) от 0 до 30° с шагом 5°;
- углы отклонения линии визирования от курса ( $\zeta$ ) от 0 до 360° с шагом 1°;
- половины углов при вершине конуса рассеяния ( $\Delta\xi$ ): 1'', 2'', 3'', 4'', 5'', 10'', 20'', 30'', 40'', 50'', 1';
- углы отклонения вспомогательной линии от курса ( $\Delta\zeta$ ) от 0 до 360° с шагом 1°.

Некоторые результаты оценки влияния точности определения УЭВО на ТОКО приведены на рис. 3-5.

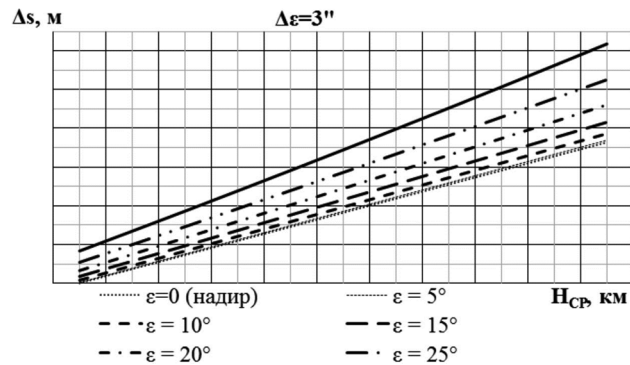


Рис. 3. Зависимость ТОКО от угла отклонения ЦЛВ от надира и средней высоты рабочей орбиты для точности определения УЭВО в 3''

На рис. 3 показано, что зависимость ТОКО для высоких точностей определения УЭВО носит линейный характер, угол наклона прямой к оси абсцисс определяется  $\varepsilon$  – углом отклонения ЦЛВ от надира. При малых углах отклонения от надира (от 0 до 10°) разница в ТОКО измеряется в единицах сантиметров вне зависи-

мости от средней высоты рабочей орбиты. По мере возрастания  $\varepsilon$  (от 10 до 30°) разница в ТОКО достигает десятков сантиметров для орбит со средней высотой от 300 до 450 км, а для орбит со средней высотой 500 км и более – единиц метров.

Точность определения УЭВО в 3'' была подтверждена для приборов БОКЗ-М, входящих в состав системы управления движением КА «Ресурс-П» как при наземной экспериментальной отработке [6], так и при его летной эксплуатации. Как видно из рис. 3, для такой точности определения УЭВО и углах отклонения от надира не более 25° значение ТОКО не превышает величины 8 м и позволяет выполнить предъявленные к КА «Ресурс-П» требования в части точности привязки снимков.

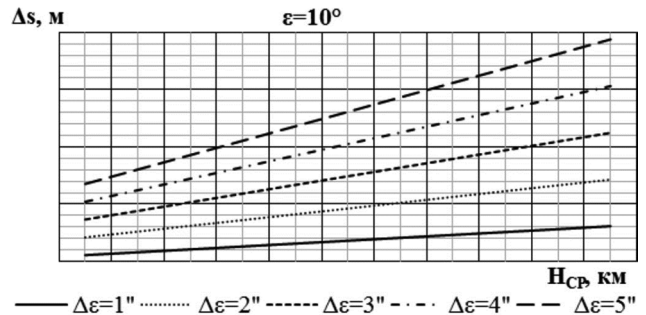


Рис. 4. Зависимость ТОКО от точности определения УЭВО и средней высоты рабочей орбиты для угла отклонения ЦЛВ от надира 10°

Как видно из рис. 4, зависимость ТОКО для малых углов отклонения ЦЛВ от надира также носит линейный характер, а угол наклона прямой к оси абсцисс определяется  $\Delta\varepsilon$  – точностью оценки УЭВО. При высоких точностях определения УЭВО (на уровне единиц угловых секунд) разница в ТОКО пропорциональна величине  $\Delta\varepsilon$  и при увеличении точности определения УЭВО на 1'' изменяется в диапазоне от 1,5 м для рабочих орбит со средней высотой 300 км и до 4 м для рабочих орбит со средней высотой 800 км. Сравнивая рис. 3 и 4, можно отметить, что  $\Delta\varepsilon$  вносит больший вклад в значение ТОКО ( $\Delta s$ ) по сравнению с углом  $\varepsilon$ .

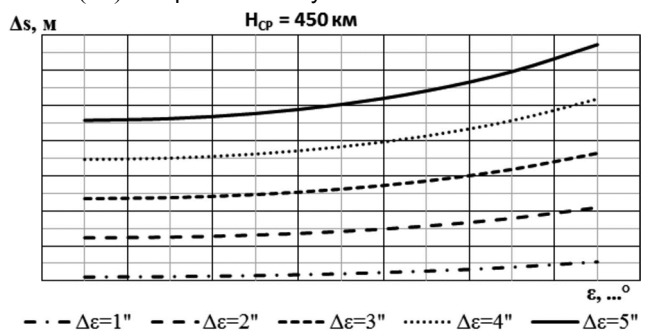
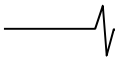


Рис. 5. Зависимость точности определения координат объекта от точности определения УЭВО и угла отклонения ЦЛВ от надира для рабочей орбиты со средней высотой 450 км

Приведенная на рис. 5 зависимость показывает наличие нелинейности в изменении ТОКО при увеличении угла  $\varepsilon$ . Причем, чем больше значение  $\Delta\varepsilon$ , тем раньше (при меньших  $\varepsilon$ ) начинает проявляться нелинейность. Так, для значений  $\Delta s$  не более 5 метров, требуется точность определения УЭВО на уровне 1-2'', при-



чем при  $\Delta\varepsilon = 2''$  угол  $\varepsilon$  должен находиться в диапазоне от 0 до  $17,5^\circ$ .

### **Заключение**

Таким образом, в работе предложен алгоритм оценки влияния точности оценки угловых элементов внешнего ориентирования КА на точность определения координат объекта, находящегося на пересечении центральной линии визирования с общеземным эллипсоидом. Представлена математическая модель предложенного алгоритма, проведена оценка влияния точности определения угловых элементов внешнего ориентирования КА на точность определения координат объекта. Результаты, приведенные в данной статье, подтверждаются данными летной эксплуатации КА «Ресурс-П» и могут быть использованы при проектировании перспективных КА оптико-электронного наблюдения.

### **Литература**

1. Кирилин А.Н., Аншаков Г.П., Ахметов Р.Н., Сторож А.Д. Космическое аппаратостроение: научно-техни-

ческие исследования и практические разработки ГНПРКЦ «ЦСКБ-Прогресс» / Под редакцией А.Н. Кирилина. – Самара: Издательский дом «АГНИ», – 2011. – 280 с.

2. Самойлов С.Ю. Модель определения географических координат объектов по космическим снимкам при помощи аналитического метода // Вестник НПО им. С.А. Лавочкина, 2011. № 4. С. 58-65.

3. Основы теории полёта космических аппаратов / Под редакцией Г.С. Нариманова и М.К. Тихонравова. – М.: Машиностроение, 1976.– 608 с.

4. Параметры общего земного эллипсоида и гравитационного поля Земли (Параметры Земли 1990 года). – М: Ред.-изд. отдел ТС ВС РФ, – 2014. – 52 с.

5. Морозов В.П. Курс сфероидической геодезии – М.:Недра, 1979.–296 с.

6. Никитин А.В., Дунаев Б.С., Красиков В.А. Анализ функционирования трех приборов звездной ориентации БОКЗ-М при съемке звездного неба // Современные проблемы дистанционного зондирования Земли из космоса. 2013. Т. 10. № 4. С. 34-42.

УДК 004.93

## ПОВЫШЕНИЕ БЫСТРОДЕЙСТВИЯ АЛГОРИТМА ФОРМИРОВАНИЯ ПАНОРАМНОГО ВИДЕОИЗОБРАЖЕНИЯ В ОПТИКО-ЭЛЕКТРОННЫХ СИСТЕМАХ С РАСПРЕДЕЛЕННОЙ АПЕРТУРОЙ

*Кудинов И.А., аспирант кафедры электронно-вычислительных машин Рязанского государственного радиотехнического университета имени В.Ф. Уткина, e-mail: igor.kudinov@mail.ru;*

*Холопов И.С., к.т.н., декан факультета радиотехники и телекоммуникаций Рязанского государственного радиотехнического университета имени В.Ф. Уткина, e-mail: kholopov.i.s@rsreu.ru.*

### IMPROVING THE PERFORMANCE OF THE ALGORITHM FOR GENERATING A PANORAMIC VIDEO IN VISION SYSTEMS WITH A DISTRIBUTED APERTURE

*Kudinov I.A., Kholopov I.S.*

*The optimization of the calculations performed in panorama vision systems during the personal region of interest (RoI) forming, in particular – bilinear interpolation, morphological dilation and spatial filtering – is considered. Analytical expressions and algorithms for reducing the number of arithmetic operations involved in the implementation of these procedures are given. It is shown that when implementing computations on single core PC processor, the use of the considered optimization algorithms in a distributed panoramic system with five cameras allows to increase the speed of RoI forming by 5,4 times.*

**Key words:** panorama vision systems, bilinear interpolation, morphological dilation, optimization algorithms.

**Ключевые слова:** панорамное изображение, матрица гомографии, блендинг, дилатация, метрика Манхэттена, билинейная интерполяция.

#### Введение

Анализ публикаций в отечественной [1-7] и зарубежной [8-12] периодике показывает актуальность исследований по формированию панорамных фото- и видеоизображений по информации от нескольких камер с пересекающимися полями зрения. Сшивка изображений с различных камер в единый панорамный кадр предполагает оценку матрицы гомографии [13-17], определяющей связь между системами координат  $i$ -й и  $j$ -й камер. Указанная оценка может выполняться по двум стратегиям: либо путем автоматического выделения особых (ключевых) точек с применением математического аппарата их дескрипторов [18-20], либо по результатам предварительной калибровки камер [21]. Первая стратегия является наиболее универсальной, но не обеспечивает хороших результатов сшивки в условиях плохой видимости, сцен с однородными текстурами, а также при объединении изображений различных спектральных диапазонов (видимого и инфракрасных). Вторая стратегия робастна к условиям съемки, однако обязательным условием ее применения является фиксированное взаимное расположение камер и наличие дальномерной информации до объектов наблюдаемой сцены.

Вне зависимости от выбранной стратегии формирование панорамы субъективно высокого качества («бесшовной» панорамы) предполагает выполнение двух операций с высокой асимптотической вычислительной сложностью (ABC):

*Рассмотрены вопросы оптимизации вычислений, выполняемых при формировании персональной области интереса (ОИ), в частности – билинейной интерполяции, морфологической дилатации и пространственной фильтрации. Приведены аналитические выражения и алгоритмы для сокращения количества арифметических операций, затрачиваемых на реализацию перечисленных процедур. Показано, что при реализации вычислений на одном ядре процессора ПЭВМ применение рассмотренных алгоритмов оптимизации в распределенной панорамной системе с пятью камерами позволяет повысить скорость формирования ОИ в 5,4 раза.*

- 1) билинейной интерполяции при поиске соответствия между пиксельными координатами камер;
- 2) сглаживания перепадов яркости (блендинга) на границах составных частей панорамного кадра, обусловленных различными условиями экспонирования при съемке с различных ракурсов.

**Целью работы** является оптимизация вычислительных затрат на реализацию билинейной интерполяции и блендинга при формировании панорамных изображений по информации от камер оптико-электронной системы с распределенной апертурой.

При оптимизации будем полагать, что все камеры формируют цифровое полутоновое изображение в формате 8 бит на цвет.

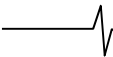
#### Варианты реализации процедуры билинейной интерполяции

Переход из системы координат камеры  $i$  в систему координат камеры  $j$ , реализуемый умножением однородных координат  $i$ -й камеры  $\mathbf{x}_i$  на матрицу гомографии  $\mathbf{H}_{ji}$ ,

$$\mathbf{x}_j = \mathbf{H}_{ji} \mathbf{x}_i, \quad (1)$$

где  $\mathbf{x}_i = [x_i, y_i, 1]^T$ ,  $\mathbf{x}_j = [x_j, y_j, 1]^T$ , а  $(x_i, y_i)$  и  $(x_j, y_j)$  – пиксельные координаты  $i$ -й и  $j$ -й камер, приводит к дробным





однородным координатам  $x_j$  [13, 22]. Поскольку значения яркости цифровых изображений определены только в узлах дискретной решетки с целочисленными пиксельными координатами [22, 23], то для определения яркости  $I(u_j, v_j)$  в точке  $(u_j, v_j)$  применяют двумерную интерполяцию. Широкое распространение в практике цифровой обработки изображений получила билинейная интерполяция (БИ) [22], суть которой сводится к двухэтапной последовательной интерполяции яркости: например, сначала по оси  $x$  – между точками  $(\lfloor x_j \rfloor, \lfloor y_j \rfloor)$ ,  $(\lfloor x_j \rfloor, \lfloor y_j \rfloor + 1)$  и  $(\lfloor x_j \rfloor + 1, \lfloor y_j \rfloor)$ ,  $(\lfloor x_j \rfloor + 1, \lfloor y_j \rfloor + 1)$ , где символом  $\lfloor \cdot \rfloor$  обозначено округление до ближайшего меньшего целого числа, а затем – по оси  $y$ . Результат БИ не зависит от порядка выполнения интерполяции по строкам и столбцам [23].

Геометрическая постановка задачи БИ [24-26] приведена на рис. 1.

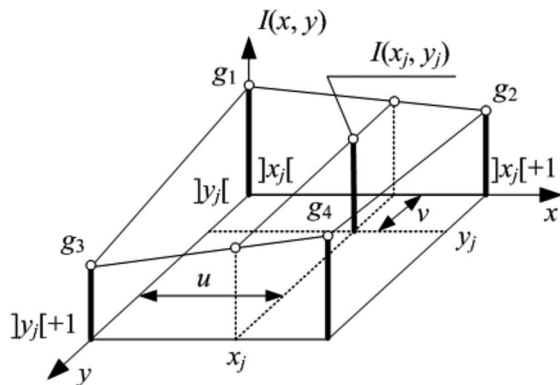


Рис. 1. Иллюстрация геометрической постановки задачи

Введя обозначения  $x = \lfloor x_j \rfloor$ ,  $y = \lfloor y_j \rfloor$ ,  $u = x_j - x$ ,  $v = y_j - y$ , получив ее решение при последовательном (двухэтапном) выполнении интерполяции по строкам и столбцам:

$$I_{ij} = I_{21} + (I_{34} - I_{21})v, \quad (2)$$

где

$$I_{21} = g_1 + (g_2 - g_1)u, \quad I_{34} = g_3 + (g_4 - g_3)u. \quad (3)$$

На рис. 1 и в формулах (2) и (3) для компактности записи приняты следующие обозначения:

$$g_1 = I(x, y), \quad g_2 = I(x+1, y), \quad g_3 = I(x, y+1), \quad g_4 = I(x+1, y+1).$$

Реализация вычисления БИ по (2) требует наличия двух дополнительных ячеек оперативной памяти для хранения интерполированных яркостей  $I_{21}$  и  $I_{34}$  помимо вспомогательных переменных  $u$  и  $v$ .

Подставив (3) в (2) и приведя подобные члены относительно значений яркости  $g_k$ ,  $k = 1..4$ , в узлах дискретной решетки, получим [25]:

$$I(x_j, y_j) = (1-x)(1-y)g_1 + x(1-y)g_2 + (1-x)yg_3 + xyg_4, \quad (4)$$

Недостатком реализации БИ непосредственно по (3) является относительно высокая вычислительная сложность: для вычисления яркости каждого пикселя требуется выполнить 8 умножений и 7 сложений.

В [25] рассмотрен способ уменьшения вычислительной сложности, согласно которому в (4) приводятся подобные члены относительно  $u$  и  $v$ :

$$I(x_j, y_j) = uv(g_4 - g_3 - g_2 + g_1) + u(g_2 - g_1) + v(g_3 - g_1) + g_1. \quad (5)$$

Как видно из (5), для выполнения БИ для каждого

пикселя требуется 4 операции умножения и 8 операций сложения.

Дополнительно уменьшить число сложений на единицу можно введением вспомогательной суммы  $g_{12} = g_1 - g_2$  или  $g_{13} = g_1 - g_3$ :

$$I(x_j, y_j) = uv(g_4 - g_3 + g_{12}) - ug_{12} + v(g_3 - g_1) + g_1 = uv(g_4 - g_2 + g_{13}) + u(g_2 - g_1) - g_{13}v + g_1. \quad (6)$$

В табл. 1 приведено сравнение времени выполнения БИ для изображения размером  $W \times H$ .

Таблица 1

Формулы для БИ	Умножений	Сложений	Количество записей вспомогательных переменных в оперативную память
(2)	$3WH$	$6WH$	$4WH$
(4)	$8WH$	$4WH$	$2WH$
(5)	$4WH$	$8WH$	$2WH$
(6)	$4WH$	$7WH$	$3WH$

Как следует из табл. 1, при выборе варианта реализации БИ следует учитывать время выполнения операций сложения, умножения и доступа к оперативной памяти для выбранного типа вычислителя. В то же время для некоторых вычислителей, например, ПЭВМ с процессором Intel Core-*i5* и оперативной памятью типа DDR4 время выполнения БИ по формулам (2), (4), (5) и (6) практически одинаково и отличается не более чем на 1 %.

### Повышение быстродействия алгоритма блендинга

При формировании панорамы блендинг является одной из наиболее вычислительно ёмких операций [8, 27, 28] и имеет АВС, стремящуюся к  $O(n^4)$ . Это объясняется тем, что блендинг требует оценивания средней яркости в окрестности каждого из пикселей, составляющих панорамное изображение.

В процедуре блендинга можно выделить несколько основных этапов [8], АВС которых без оптимизации следующая:

1) формирование карт весовых коэффициентов для формирующих панораму изображений на основе меры удаления пикселя от центра кадра – АВС  $O(n^2)$ ;

2) пространственная фильтрация изображений с камер с АВС  $O(w^2n^2)$  для оценивания низкочастотной (НЧ), среднечастотной (СЧ) и высокочастотной (ВЧ) компонент, где  $w$  – горизонтальный и вертикальный размеры апертуры пространственного фильтра, а также аналогичная фильтрация для карт весовых коэффициентов; при  $w$ , стремящихся к размерам ОИ, АВС данного этапа будет стремиться к  $O(n^4)$ .

3) коррекция яркости пикселей на основе карт весовых коэффициентов – АВС  $O(n^2)$ .

Рассмотрим методы оптимизации вычислений для второго этапа блендинга, АВС которого более  $O(n^2)$ .

### Повышение быстродействия низкочастотной фильтрации с применением ВОХ-фильтра

Для оценки яркости фоновой составляющей рационально применять процедуру быстрого сглаживания исходного изображения ВОХ-фильтром [29, 30]. Дополни-

тельно снизить вычислительные затраты позволяет использование подхода, рассмотренного в [31, 32]. Если при этом без нарушения общности задачи пространственной линейной фильтрации ввести дополнительное условие, что апертура фильтра  $w \times w = 2^p \times 2^p$ ,  $p \in \mathbb{N}$ , то для вычисления отфильтрованного значения  $Y_{\phi i}$  1D ВОХ-фильтра требуется всего 2 сложения (вычитание яркости пикселя, вышедшего из апертуры фильтра, и суммирование с яркостью пикселя, накрытого ей) и одна операция сдвига на  $p$  разрядов вправо:

$$acc = acc + Y_{i+2^{p-1}} - Y_{i-2^{p-1}}, Y_{\phi i} = acc \gg p, \quad (7)$$

где  $acc$  – значение суммы яркостей пикселей, накрытых апертурой (аккумулятор),  $\gg$  – бинарный оператор сдвига вправо. Выражение (7) справедливо, если для индекса пикселя  $i$  выполняется неравенство:

$$2^{p-1} < i < W - 2^{p-1}.$$

Для реализации процедуры фильтрации по (7) требуется расширение изображения по высоте и ширине на размер апертуры фильтра [31] – по  $2^{p-1}$  строк и столбцов сверху, снизу, слева и справа.

В силу линейности ВОХ-фильтра 2D фильтрация сводится к двум последовательным проходам по изображению 1D ВОХ-фильтром [33]: по строкам и по столбцам. В результате АВС 2D низкочастотного пространственного фильтра уменьшается с  $O(w^2 n^2)$  до  $O(w n^2)$ .

#### Повышение быстродействия алгоритма многодиапазонного выравнивания яркости

Одним из популярных алгоритмов выравнивания яркости при формировании панорамного кадра является многодиапазонный блендинг (*Multi-band blending*) [8]. Его основная идея заключается в том, что для смешения низкочастотных составляющих яркости исходных изображений применяют пространственную апертуру большего размера, чем для смешения высокочастотных составляющих [34]. Поэтому на каждом  $k$ -м шаге многодиапазонного блендинга для каждого пикселя панорамного изображения с координатами  $(x, y)$  находится взвешенная сумма яркостей НЧ  $I_{НЧ}$ , СЧ  $I_{СЧ}$  и ВЧ  $I_{ВЧ}$  компонент яркости для  $N$  формирующих панораму изображений, нормированная к сумме весовых коэффициентов:

$$I_{(k)xy}^{\text{бленд}} = \frac{\sum_{i=1}^N W_{(k)xy}^{(i)} I_{(k)xy}^{(i)}}{\sum_{i=1}^N W_{(k)xy}^{(i)}}, \quad (8)$$

где  $k = \{\text{НЧ}, \text{СЧ}, \text{ВЧ}\}$ ,  $i = 1, 2, \dots, N$ , а  $W_{(k)}^{(i)}$  и  $I_{(k)}^{(i)}$  – соответственно матрицы весов и яркостей пикселей  $k$ -го диапазона частот  $i$ -го изображения, к которому применено обратное проективное преобразование (1) для перехода от системы пиксельных координат ОИ  $(x, y)$  в систему координат  $i$ -й камеры:

$$I_{xy}^{(i)} = I(x_i, y_i),$$

где  $[x_i, y_i, 1]^T = \mathbf{H}_{\text{ОИ}i}^{-1}[x, y, 1]^T$ ,  $\mathbf{H}_{\text{ОИ}i}$  – матрица гомографии, определяющая переход от пиксельных координат  $(x_i, y_i)$  к пиксельным координатам  $(x, y)$ . Итоговое изображение формируется как сумма НЧ, СЧ и ВЧ компонент, к которым применен блендинг:

$$\mathbf{I}^{\text{бленд}} = \sum_k \mathbf{I}_{(k)}^{\text{бленд}}. \quad (9)$$

Для формирования матриц весовых коэффициентов  $W_{(k)}^{(i)}$  требуется выполнить следующие шаги [8, 34]:

1) рассчитать матрицы удаления пикселей с координатами  $(x_i, y_i)$  от центра изображения  $i$ -й камеры по горизонтали  $D_X^{(i)}$  и вертикали  $D_Y^{(i)}$ :

$$D_{Xxy}^{(i)} = |x_i - 0,5W^{(i)}|, \quad D_{Yxy}^{(i)} = |y_i - 0,5H^{(i)}|, \quad (10)$$

где  $W^{(i)}$  и  $H^{(i)}$  – соответственно ширина и высота  $i$ -го изображения;

2) сформировать матрицы весовых коэффициентов  $V^{(i)}$ , элементы которых принимают значение 1 при нулевом удалении и 0 – при максимуме:

$$V_{xy}^{(i)} = \frac{2(1 - D_{Xxy}^{(i)})}{W^{(i)}} \cdot \frac{2(1 - D_{Yxy}^{(i)})}{H^{(i)}} = \frac{4}{W^{(i)} H^{(i)}} (1 - D_{Xxy}^{(i)}) (1 - D_{Yxy}^{(i)}); \quad (11)$$

3) выполнить бинаризацию элементов матриц  $V_{\text{бин}}^{(i)}$ :

$$V_{\text{бин}_xy}^{(i)} = \begin{cases} 1, & \text{если } V_{xy}^{(i)} = \max_{i=1..N} (V_{xy}^{(i)}), \\ 0, & \text{в противном случае.} \end{cases} \quad (12)$$

4) выполнить свертку матриц бинарных элементов  $V_{\text{бин}}^{(i)}$  с маской пространственного НЧ фильтра  $g^{(k)}$  размерности  $w^{(k)} \times w^{(k)}$  (в [8, 29] применяют гауссовский фильтр):

$$W_{(k)}^{(i)} = V_{\text{бин}}^{(i)} \otimes g^{(k)}, \quad (13)$$

где  $\otimes$  – оператор свертки. Для дополнительного сокращения вычислений при оценке компонент  $I_{(k)}^{(i)}$  вместо гауссовского фильтра целесообразно применять ВОХ-фильтр. При этом (13) преобразуется к виду:

$$W_{(k)}^{(i)} = V_{\text{бин}}^{(i)} \otimes \mathbf{A}^{\text{ВОХ}(k)},$$

где  $\mathbf{A}^{\text{ВОХ}(k)}$  – маска ВОХ-фильтра для  $k$ -го диапазона пространственных частот.

Для распределенных панорамных систем, камеры которых жестко закреплены и в процессе эксплуатации не перемещаются друг относительно друга, вычисления по формулам (10)-(13) выполняются однократно – только на этапе инициализации [21].

При относительно больших угловых размерах пересечения полей зрения камер панорамной системы (10 % и более от ширины поля зрения каждой камеры) для сокращения АВС можно выполнять блендинг по (7) не для всех пикселей изображения, а только в пределах некоторой области – вблизи границы пересечения бинарных изображений  $V_{\text{бин}}^{(i)}$ . Такое упрощение допустимо при ширине данной области не менее  $0,1 W$ . При этом выигрыш в производительности будет тем больше, чем больше угловые размеры пересечения полей зрения камер по сравнению с угловыми размерами области выполнения блендинга.

Для определения области применения блендинга в ОИ необходимо:

1) сформировать бинарную маску (БМ)  $m_{\text{бин}}$  с границами пересечения в ОИ изображений  $V_{\text{бин}}^{(i)}$ , составляющих панорамное поле – АВС  $O(n^2)$ ;

2) расширить область БМ с  $m_{\text{бин}}$  до  $M_{\text{бин}}$ , выполнив морфологическую операцию дилатации (сложения Мин-

```

// первый проход
for (j = 0; j < H; j++)
  for (i = 0; i < W; i++) {
    if (Yij == 1)
      dij = 0;
    else {
      dij = W + H;
      if (j > 0)
        dij = min(dij, di(j-1) + 1);
      if (i > 0)
        dij = min(dij, d(i-1)j + 1);
    }
  }

```

```

// второй проход
for (j = H-1; j ≥ 0; j--)
  for (i = W-1; i ≥ 0; i--) {
    if (j+1 < H)
      dij = min(dij, di(j+1) + 1);
    if (i+1 < W)
      dij = min(dij, d(i+1)j + 1);
  }
// вычисление дилатации
for (j = 0; j < H; j++)
  for (i = 0; i < W; i++)
    if (dij < 0,5D - 1)
      Yдилij = 1;
    else
      Yдилij = 0;

```

Рис. 2. Псевдокод процедуры вычисления дилатации через МР

ковского) [35] с апертурой структурирующего элемента  $D \times D$  и АВС  $O(D^2 n^2)$ .

Как видно, такое упрощение хотя и позволяет полностью отказаться от блендинга на участках ОИ, не накрытых БМ  $\mathbf{M}_{\text{блн}}$ , но в то же время увеличивает АВС за счет необходимости выполнения процедуры морфологической дилатации.

### Повышение быстродействия алгоритма дилатации

Оптимизация операции морфологической дилатации бинарного изображения с апертурой структурирующего элемента  $D \times D$  приводится в [36] и заключается в следующем. Для нулевых пикселей бинарного изображения строится карта расстояний  $\mathbf{d} = \{d_{ij}\}$ ,  $i = 1..W, j = 1..H$ , до ближайших к ним единичных пикселей с координатами  $(x_{i1}, y_{i1})$ , причем мерой расстояния  $d$  выступает метрика Манхэттена (манхэттенское расстояние,  $L_1$ -норма Минковского, расстояние городских кварталов, метрика такси) [37, 38]:

$$d_{ij} = L_1[(x_i, y_i), (x_{i1}, y_{i1})] = |x_i - x_{i1}| + |y_i - y_{i1}|.$$

При этом морфологическая операция дилатации сводится к сравнению манхэттенского расстояния (МР) для каждого пикселя с порогом  $0,5D - 1$ :

$$I_{\text{дил<sub>ij</sub>} = \begin{cases} 1, & d_{ij} \leq 0,5D - 1, \\ 0, & \text{в противном случае.} \end{cases}$$

Для формирования карты МР требуется всего два прохода по изображению [36]:

1) от левого верхнего угла (ЛВУ) к нижнему правому углу (НПУ), в ходе которого рассчитываются МР при удалении от пикселя в направлениях «вниз» и «вправо»;

2) в обратной последовательности – от НПУ к ЛВУ, в ходе которого рассчитываются МР при удалении от пикселя в направлениях «вверх» и «влево».

Псевдокод, реализующий оптимизированный алгоритм дилатации с апертурой структурирующего элемента  $k \times k$  на основе карты МР, приведен на рис. 2.

Код программы на языке C++, реализующий псевдокод рис. 2, приводится в [36].

Следует отметить, что рассмотренные методы повышения быстродействия не применимы при реализации параллельных вычислений на графических ускорителях (например, поддерживающих технологию CUDA) или многоядерных процессорах. Это связано с тем, что обработка для каждого пикселя изображения согласно

(7) и процедуре вычисления дилатации через МР не является однородной [22].

**Результаты полунатурного эксперимента.** Исходные изображения для эксперимента получены с макета панорамной системы (рис. 3) [21, 39], содержащего управляемую пользователем персональную область интереса (ОИ).



Рис. 3. Макет мультиспектральной распределенной панорамной системы: сверху – пять телевизионных камер, снизу по центру – камера длинноволнового инфракрасного (ИК) диапазона 8-14 мкм, в левом нижнем углу – камера коротковолнового ИК диапазона 0,9-1,7 мкм

На рис. 4 приведена ОИ размера  $1024 \times 768$  пикселей и шириной поля зрения  $40^\circ \times 30^\circ$ , содержащая сектора с изображениями от трех камер макета, а на рис. 5 и 6 – результаты промежуточных вычислений для реализации упрощенного блендинга: границы изображений камер в ОИ (рис. 5) и бинарная маска  $\mathbf{M}_{\text{блн}}$ , по которой выполняется сглаживание яркости (рис. 6). Изображение рис. 6 получено в результате применения к рис. 5 морфологической дилатации с размером структурного элемента  $D \times D = 127 \times 127$  пикселей согласно псевдокоду рис. 2, а размеры апертуры ВОХ-фильтра, принятые в работе для реализации многодиапазонного блендинга, составили соответственно  $w_{\text{нч}} \times w_{\text{нч}} = 128 \times 128$ ,  $w_{\text{сч}} \times w_{\text{сч}} = 16 \times 16$  и  $w_{\text{вч}} \times w_{\text{вч}} = 4 \times 4$  пикселей.

Темп обновления информации при формировании ОИ по информации от пяти телевизионных камер и размерности ОИ  $1024 \times 768$  на ПЭВМ с процессором Intel Core i5 без процедур оптимизации, рассмотренных в работе, и с ними приведен в табл. 2. Как было отмечено

выше, для данного типа вычислителя все стратегии выполнения билинейной интерполяции обеспечивают приблизительно одинаковое быстродействие.



Рис. 4. ОИ без блендинга

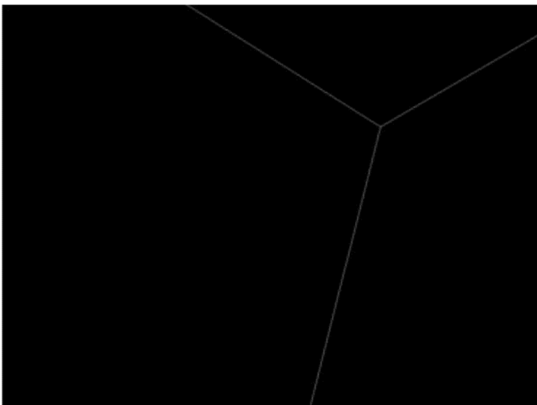


Рис. 5. Границы изображений камер в ОИ (БМ  $m_{\text{бин}}$ )

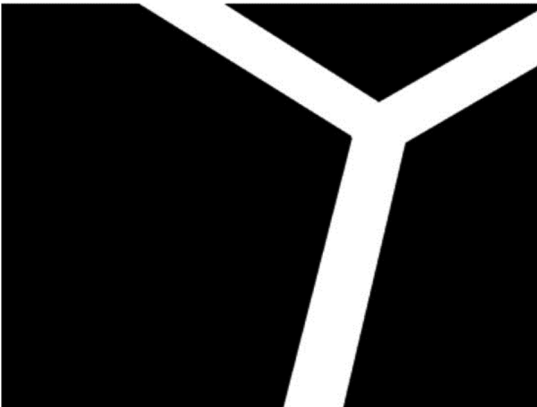


Рис. 6. БМ  $M_{\text{бин}}$ , полученная в результате дилатации



Рис. 7. ОИ после блендинга

Таблица 2

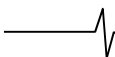
	Без оптимизации	С оптимизацией
Темп обновления информации в ОИ, Гц	1,24	6,7

### Заключение

Таким образом, из полученных в ходе макетирования численных данных следует, что приведенные процедуры оптимизации и упрощения вычислений для реализации пространственной фильтрации, морфологической дилатации и блендинга позволяют повысить темп формирования ОИ более чем в 5 раз.

### Литература

1. Лазаренко В.П., Джамийков Т.С., Коротаев В.В., Ярышев С.Н. Метод создания сферических панорам из изображений, полученных всенаправленными оптико-электронными системами // Научно-технический вестник информационных технологий, механики и оптики. 2016. Т. 16, № 1. С. 46-53.
2. Прудников Н.В., Шлишевский В.Б. Панорамные оптико-электронные устройства кругового и секторного обзора // Вестник СГУГиТ. 2016. Т. 33, Вып. 1. С. 148-161.
3. Бельский А., Жосан Н., Брондз Д., Горбачев К., Гребенчиков В., Каргаев А. Круглосуточная панорамная система технического зрения для вертолетов // Фотоника. 2013. Т. 38, №2. С. 80-86.
4. Багдасарова О.В., Багдасаров А.А. Панорамные системы кругового обзора в устройствах отображения вторичной информации комплексов авионики и автобазиирования // Фазотрон. 2012. № 3 (19). С. 28-33.
5. Аршакян А.А. Панорамное наблюдение сцены с борта летательного аппарата // Известия ТулГУ. Технические науки. 2013. Вып. 4. С. 144-151.
6. Широков Р.И., Алехнович В.И. Увеличение углов поля зрения составной веб-камеры методом склеивания изображений // Контентант. 2014. № 4. С. 10-23.
7. Сильвестрова О.В. Оценка эффективности способов цветовой коррекции для панорамных изображений с малоразмерными объектами // Цифровая обработка сигналов. 2015. № 3. С. 25-28.
8. Brown M, Lowe D. Automatic Panoramic Image Stitching using Invariant Features // International Journal of Computer Vision. 2007. Vol. 74(1). P. 59-73.
9. Szeliski R. Image alignment and stitching: a tutorial // Foundations and trends in computer graphics and vision. 2006. Vol. 2(1). P. 1-104.
10. Perazzi F., Sorkine-Hornung A., Zimmer H., Kaufmann P., Wang O., Watson S., Gross M. Panoramic video from unstructured camera arrays // Comput. Graph. Forum. 2015. Vol. 34(2). P. 57-68.
11. Agarwala A., Zheng K.C., Pal C., Agrawala M., Cohen M., Curless B., Salesin D., Szeliski R. Panoramic video textures // ACM Transactions on Graphics. Vol. 24(3). P. 821-827.
12. Shum H.-Y., Szeliski R. Construction of panoramic mosaics with global and local alignment // International J. of Computer Vision. 2000. Vol. 36(2). P. 101-130.



13. Hartley R., Zisserman A. Multiple view geometry in computer vision: 2nd edition. Cambridge: Cambridge University Press, 2003. 656 p.
14. Faugeras O., Lustman F. Motion and structure from motion in a piecewise planar environment // *International J. of Pattern Recognition and Artificial Intelligence*. 1988. Vol. 2(3). P. 485-508.
15. Malis E., Vargas M., Deeper understanding of the homography decomposition for vision-based control, Technical report, INRIA, 2007.
16. Ефимов А.И., Новиков А.И. Алгоритм поэтапного уточнения проективного преобразования для совмещения изображений // *Компьютерная оптика*. 2016. Т. 40, № 2. С. 258-265.
17. Стротов В.В. Сравнение двух структурных алгоритмов оценки параметров геометрических преобразований изображений // *Цифровая обработка сигналов*. 2013. № 3. С. 21-25.
18. Lowe D.G. Distinctive Image Features from Scale-Invariant Keypoints // *International Journal of Computer Vision*. 2004. Vol. 60(2). P. 91-110.
19. Fischler M., Bolles R. Random Sample Consensus: a Paradigm for Model Fitting with Application to Image Analysis and Automated Cartography // *Communications of the ACM*. 1981. Vol. 24(6). P. 381-395.
20. Novikov A.I., Sablina V.A., Nikiforov M.B., Logvinov A.A. The contour analysis and image-superimposition problem in computer vision systems // *Pattern Recognition and Image Analysis*. 2015. Vol. 25(1). P. 73-80.
21. Kudinov I.A., Pavlov O.V., Kholopov I.S., Khramov M.Yu. The algorithm for a video panorama construction and its software implementation using CUDA technology // *CEUR Workshop Proceedings*. Vol. 1902. 2017. P. 37-42.
22. Грузман И.С., Киричук В.С., Косых В.П., Перетягин Г.И., Спектор А.А. Цифровая обработка изображений в информационных системах: учеб. пособие. Новосибирск: Изд-во НГТУ, 2002. 352 с.
23. Красильников Н.Н. Цифровая обработка 2D- и 3D-изображений: учеб. пособие. СПб.: БХВ-Петербург. 2011. 608 с.
24. Parker J.A., Kenyon R.V., Troxel D.E. Comparison of interpolating methods for image resampling // *IEEE Trans. on Medical Imaging*. 1983. Vol. 2(1). P. 31-39.
25. Li J., Su J., Zeng X. A solution method for image distortion correction model based on bilinear interpolation // *Computer Optics*. 2019. Vol. 43(1). P. 99-104.
26. Тимофеева Н.Е., Гераськин А.С. Исследование возможности улучшения алгоритма билинейной интерполяции для корректировки цифровых изображений применением теории полей ориентации // *Вестник ВГУ. Серия: системный анализ и информационные технологии*. 2018. № 1. С. 119-125.
27. Zhu Z., Lu J., Wang M., Zhang S., Martin R.R., Liu H., Hu S.-M. A comparative study of algorithms for realtime panoramic video blending // *IEEE Trans. on image processing*. 2018. Vol. 27, No. 6. P. 2952-2965.
28. Канаева И.А., Болотова Ю.А. Методы коррекции цвета и яркости при создании панорамных изображений // *Компьютерная оптика*. 2018. Т. 42, № 5. С. 885-897.
29. Jähne B. Digital image processing: 6th ed, revised and extended. New York: Springer, 2005. 585 p.
30. Муратов Е.Р., Устюков Д.И. Алгоритм выравнивания освещённости объектов на изображениях с поточным выполнением // *Техническое зрение в системах управления - 2017: тезисы докладов*. М.: ИКИ РАН, 2017. С. 29-31.
31. Lukin A. Tips & Tricks: Fast Image Filtering Algorithms // *GraphiCon: Proc. of GraphiCon'2007Russia conference, Moscow, June 23-27, 2007*.
32. Jarosz W. Fast Image Convolutions [Электронный ресурс]. Режим доступа: [http://elynxsdk.free.fr/ext-docs/Blur/Fast\\_box\\_blur.pdf](http://elynxsdk.free.fr/ext-docs/Blur/Fast_box_blur.pdf). Дата обращения: 17.05.2017.
33. Pratt W.K. Digital image processing: 4th ed. New Jersey: John Wiley & Sons, 2007. 807 p.
34. Burt P., Adelson E. A multiresolution spline with application to image mosaics // *ACM Transactions on Graphics*. 1983. Vol. 2(4). P. 217-236.
35. 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.
36. Efficiently implementing dilate and erode image functions [Электронный ресурс]. Режим доступа: <https://blog.ostermiller.org/dilate-and-erode> (дата обращения: 17.10.2017).
37. Minkowski H. *Geometrie der zahlen*. Leipzig und Berlin: Druck und Verlag von B.G. Teubner, 1910. 256 p.
38. Krause E.F. *Taxicab geometry: an adventure in non-euclidean geometry*. New York: Dover Publications, 2012. 96 p.
39. Кудинов И.А., Холопов И.С., Храмов М.Ю. Технология формирования панорамных разноточных видеозаписей для обзорных авиационных оптико-электронных систем // *Труды МАИ*. 2019. Вып. 104. С. 1-20.

УДК 520.34, 520.6

## УВЕЛИЧЕНИЕ ДИНАМИЧЕСКОГО ДИАПАЗОНА ВИДЕОСИСТЕМЫ ЛОГИЧЕСКИМ СЛОЖЕНИЕМ ЦИФРОВЫХ ИЗОБРАЖЕНИЙ

*Котцов В.А., научный сотрудник Института космических исследований Российской академии наук (ИКИ РАН), e-mail: vladkott@mail.ru*

### INCREASE THE DYNAMIC RANGE OF THE VIDEO SYSTEM LOGICAL ADDITION OF DIGITAL IMAGES

*Kottsov V.A.*

*The possibility of rapid increase of the dynamic range of digital images by obtaining images with mutually additional characteristics and their parallel logical summation in the streaming mode is shown.*

**Key words:** dynamic range, digital images, logical summation, streaming mode.

**Ключевые слова:** динамический диапазон, многоканальное наблюдение, логическое сложение, потоковая обработка.

*Показана возможность оперативного увеличения динамического диапазона цифровых изображений путем получения изображений с взаимно дополнительными характеристиками и их параллельного логического суммирования в потоковом режиме.*

#### Введение

Видеоинформационные системы играют важную роль в изучении природы Земли и других планет, в навигации, обеспечении безопасности, робототехнике и других важных приложениях. Материалы аэро и космической съемки, в частности, служат важным источником информации для изучения земных ресурсов, состояния природной среды, условий существования на других планетах и многих других направлений исследований необходимых для человеческой деятельности. Результаты этих исследований прямо зависят от качества получаемой видеоинформации. Однако условия наблюдения с космических аппаратов на разных участках орбиты постоянно меняются. На поверхности Луны и других небесных тел, не имеющих атмосферы, условия наблюдений характеризуются резкими изменениями освещенности различных участков. Телевизионные средства наружного наблюдения в системах обеспечения безопасности тоже вынуждены работать в широком диапазоне изменения условий освещения.

#### Роль динамического диапазона для видеосистемы

Одной из основных характеристик съемочной системы является ее динамический диапазон, который определяет возможность отображения всего многообразия яркостей объектов, наблюдаемых в поле зрения. Большой диапазон изменения яркости объектов реальных сцен далеко не всегда вписывается в ограниченные технические возможности средств наблюдения. При этом может теряться часть необходимой видеоинформации.

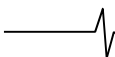
При наблюдении сцен в естественных условиях освещенность земной поверхности в течение суток меняется в широком диапазоне. В полдень ясного дня освещенность максимальна и доходит до  $10 \times 10^4$  лк. Диапазон изменений освещенности с течением суток составляет около 180 Дб. Для адаптации съемочных систем к этим изменениям применяют меняющееся

диафрагмирование входной оптической системы, светофильтры с различным пропусканием, выбирают соответствующую освещенности светочувствительность фотоприемника, а также время его экспонирования (накопления сигнала).

Различные участки сцены могут иметь существенно разные условия освещения. Это приводит к тому, что в реальных условиях наблюдения, часть яркостной информации может выходить за рамки области линейного отображения и, следовательно, не будет правильно отображаться на снимке. Всем хорошо известны потери деталей на фотографиях в «теньях» или «светах». Например, вы фотографируете в неосвещенной комнате и за окном солнечный день. Глазами хорошо видны предметы, как в комнате, так и за окном. Но на снимке, в зависимости от параметров съемки, мы увидим либо предметы в комнате, либо за окном. Типичным примером может служить различие условий фотографирования на уровне крон деревьев и под пологом леса при аэрофотосъемке, которое может быть критичным для фотограмметрических определений. Наличие облачности при съемке тоже приводит к несоответствию условий отображения на снимке объектов с прямым и рассеянным освещением.

Еще одним важным примером высоких требований к динамическому диапазону служит задача наблюдения полного лимба планеты в произвольной фазе, например, для навигационного определения направления местной вертикали.

Камера с большим динамическим диапазоном иногда позволяет обеспечить получение видеоинформации даже в условиях далеких от идеальных. Так цифровая камера в составе научного комплекса микроспутника «Чибис-М» за счет большого динамического диапазона обеспечила получение ценной видеоинформации даже в режиме неуправляемой съемки. Были получены мелкомасштабные снимки разных регионов, снимки горизонта с проработкой атмосферных слоев, ночные снимки с



населенными пунктами, которые позволили дать оценки скорости и направления перемещения спутника [1]. Нестандартную задачу наблюдения звезд в дневных условиях равнинной местности решали специалисты из Сарова. По результатам сравнительных испытаний образцов камер зарубежного и российского производства, уверенные дневные наблюдения звезд слабой яркости 7-й и 8-й звездной величины обеспечила только российская камера RT-1000DC производства ООО «Растр Технолоджи» [2].

Эти несколько примеров показывают важность обеспечения необходимого динамического диапазона системы наблюдения при решении конкретных задач.

### **Проблема увеличения динамического диапазона**

Чтобы получить всё желаемое отображение сцены нужно выбрать такие параметры съемки, чтобы весь необходимый диапазон градаций наблюдаемой яркости укладывался в пределах участка линейного отображения. Однако, желание увеличить диапазон отображаемой яркости нередко приводит к соответствующему снижению контраста получаемого изображения, что в свою очередь ведет к размыванию границ и уменьшению распознаваемости объектов.

Любая съемочная аппаратура характеризуется определенным ограниченным динамическим диапазоном. Он может задаваться различным образом. Для фотографических систем его определяет фотографическая широта, для цифровых систем наблюдения – число передаваемых градаций яркости. Манипулируя величиной контраста отображения и величиной экспозиции, подбирают оптимальные условия получения изображения. Однако, реальные условия наблюдения и технические возможности аппаратуры не всегда позволяют этим путем достигнуть необходимого результата. В качестве примера сложного для отображения объекта сошлемся на особенности распределения яркости при наблюдении картины оптического преобразования Фурье, в центральной части поля зрения наблюдается высокая яркость, соответствующей низким пространственным частотам, а к краям поля зрения, с увеличением пространственных частот происходит быстрое падение яркости наблюдаемой картины пространственного спектра. Правильное отображение этой картины весьма проблематично.

В настоящее время для получения видеoinформации широко используют цифровые съемочные системы с матричными фотоприемниками, которые обеспечивают получение дискретизованных и квантованных изображений. Однако, их динамический диапазон всё еще меньше чем у фотоматериалов и часто бывает недостаточен для решения многих прикладных задач.

Если результаты съемки важны и снижение требований к качеству изображения нежелательно, то следует искать путь увеличения динамического диапазона без снижения контраста формируемого изображения.

### **Использование многоканального наблюдения**

Из постановки задачи следует необходимость увеличения числа градаций получаемого изображения, но

оно всегда будет ограничено техническими возможностями конкретной съемочной системы.

Вариантом решения этой задачи может быть многократное или параллельное получение изображений одной и той сцены при разных, дополняющих друг друга параметрах съемки. Наиболее просто, например, для статической сцены можно выполнять съемку с изменением времени экспонирования при фотографировании для каждого последующего кадра, что соответствует изменению времени накопления заряда в случае съемки камерой с твердотельным матричным фотоприемником. Каждое из полученных таким образом изображений будет характеризовать только часть видеoinформации, заключенной в пределах области линейного отображения. Следовательно, каждый из полученных в серии снимков будет дополнять наблюдаемую картину и для последующей интерпретации следует рассматривать одновременно все полученные снимки. Но, выполнение такого анализа на нескольких отдельных снимках заметно усложняется и становится менее эффективным.

Современные цифровые съемочные системы могут формировать изображения с 8, 10, 12 и большим числом бит, а цифровая видеoinформация может затем поступать в процессорные блоки, обрабатываться математически и анализироваться без принципиального ограничения к числу градаций. Полученный результат будет оцениваться в виде цифрового массива или представлен для отображения в соответствии с тем числом градаций, которое может реализовать применяемая техника воспроизведения. Следовательно, основные ограничения числа градаций мы имеем на этапе получения видеoinформации.

Можно указать множество случаев, когда последовательная схема для получения набора изображений с разными свойствами неосуществима. Достаточно привести любой пример быстро изменяющейся ситуации в процессе наблюдения. В этом случае съемку такого набора изображений с разными свойствами необходимо выполнять только одновременно. Подобная технология одновременной съемки с разными параметрами известна и широко применяется на практике, в частности, для определения спектральных различий объектов с использованием многозональной (многоканальной) съемочной системы [3]. Допустим, что мы получили требуемый набор таких изображений, отображающих весь необходимый динамический диапазон. Очевидно, что для практических целей эту одновременно полученную видеoinформацию целесообразно скомпилировать в одно изображение. Так же поступают и при представлении для анализа результатов многозональной съемки, синтезируют цветокодированное изображение из отдельных кадров, полученных в разных зонах спектра.

Но, в ряде случаев существенное значение может иметь выбор технологии для выполнения этой компиляции из серии полученных при разных условиях изображений. От этого зависит выбор используемой техники, её скоростные характеристики и стоимость используемого оборудования. Это может быть важно в бортовой аппаратуре и особенно если возникает необходимость оперативного решения этой задачи в темпе поступления видеoinформации.



### Вариант арифметического решения

Как следует из общей формулировки задачи, требуемое решение может осуществляться, например, арифметическим сложением видеосигналов изображений, для соответствующих точек изображения, полученных на двух и более кадрах, а затем последующей нормализации результата, если это необходимо. Такое решение предложено, в частности в [4]. Однако, сложение цифровой видеоинформации в процессорных устройствах является не совсем простой процедурой. Необходимо многократно осуществить последовательную выборку соответствующих пикселей изображений, полученных при разных условиях экспонирования. Значения цифровых сигналов, характеризующих наблюдения в одной и той же точке сцены, необходимо просуммировать.

Всякое устройство, выполняющее действия над цифровыми сигналами, является функциональным преобразователем, в основе работы которого применяют логические элементы. Архитектура ЭВМ состоит из типовых логических устройств, работающих на основе аппарата математической логики с определением переполнения разрядов и переносом данных в более высокие порядки. Для обеспечения работы таких арифметических устройств в их схемах предусмотрены блоки запоминания и синхронизации [5]. Очевидно, что в процессе вычислений, в зависимости от количества накоплений и переносов, происходит переменная задержка и скорость обработки данных меняется. Это усложняет арифметическое решение задачи, реализуемое в устройствах электронно-вычислительной техники.

### Вариант логического решения

Нами предложен логический вариант решения задачи увеличения динамического диапазона цифровых приборов. Он также заключается в получении при съемке набора изображений одной и той сцены при разных, дополняющих друг друга параметрах съемки и последующего формирования суммарного изображения, которое содержит всю полученную видеоинформацию. Отличие заключается в технологии выполнения самой операции суммирования.

Рассмотрим этот вариант логического решения задачи. В современных твердотельных электронных съемочных системах изображение формируют в виде матрицы дискретных значений яркости, наблюдаемой в каждой разрешаемой точке. Величину самой яркости в цифровом изображении также дискретизируют или квантуют на заданное ограниченное число уровней. Любой из таких выбранных уровней яркости характеризует фотометрический срез изображения и может быть представлен, как одномерный слой информации кодированный в бинарной форме. Каждый такой слой в бинарном представлении содержит только значения 0 и 1. Восемьбитное изображение, например, можно представить квантованным набором изображений, как наложение 256 одноуровневых слоев бинарных матриц.

В соответствии с требованием попиксельного совмещения всего полученного набора изображений, представленную таким образом картину квантованных уров-

ней одного кадра мы можем дополнить квантованными уровнями, полученными в другом кадре. Сложность заключается в том, что такого простого совмещения будет недостаточно для получения правильной картины. Структура изображения имеет определенную упорядоченность информации по уровням, которая нарушается при простом наложении бинарных слоев двух разных изображений. Это происходит всегда, даже для изображений одной и той же сцены. Организация размещения квантованных слоев в любом изображении такова, что все единичные значения должны быть сосредоточены на нижних уровнях, а нулевые значения на верхних уровнях над ними. То есть, между слоями со значениями единица не должно быть слоев со значениями ноль. При простом совмещении структур изображений это условие нарушается.

Заметим, что каждый из квантованных слоев цифрового изображения можно также рассматривать, как пространственную совокупность логических сигналов (0 и 1). Если считать уровни с 1 заполненными, а уровни с 0 пустыми, то пространственное отображение многослойной модели такой видеоинформации легко представить в виде рельефа поверхности образованной наложением заполненных уровней единиц, например.

Известно также, что для двух бинарных сигналов эта же процедура арифметического сложения может быть представлена наложением двух логических операций – операции логического сложения и операции логического умножения [6]. Математическое выражение суммы, получаемой для логических операторов в этом случае можно записать следующим образом

$$\omega(\Sigma) = 1 \# (x \vee y) \# (x \& y) \# 0,$$

где  $\&$  – логическая операция И, а  $\vee$  – логическая операция ИЛИ.

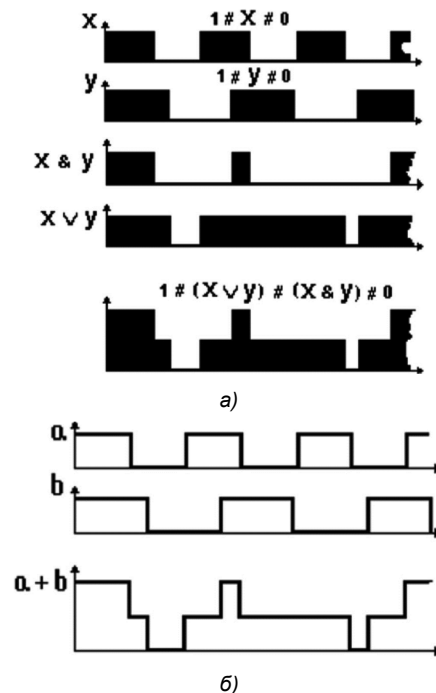


Рис.1. Сравнительные особенности выполнения операции сложения для логической процедуры (а) и для арифметической процедуры (б)

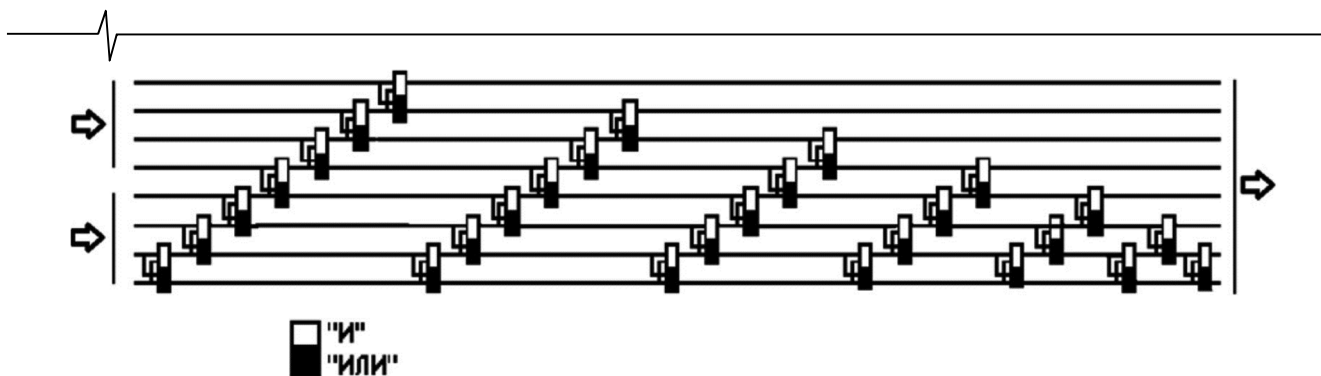


Рис. 2. Суммирующее устройство для квантованных сигналов на двойных логических элементах И и ИЛИ

На рис. 1 показаны сравнительные особенности операции сложения двух бинарных сигналов в случае выполнения логической процедуры (рис. 1 а) и в случае выполнения арифметической процедуры (рис. 1 б). Опираясь на сказанное, можно организовать схему, в которой арифметические операции заменяются логическими операциями между значениями на квантованных уровнях для выполнения совмещения видеoinформация двух или более кадров.

Однако, размещение видеoinформации по уровням при выполнении этих операций для изображений с несколькими уровнями квантования нарушается. Для восстановления правильного размещения видеoinформации по уровням следует все значения с 1 опустить, а значения с 0 поднять. То есть, для формирования изображения необходимо вернуться к принятой нами модели размещения логических сигналов по уровням путем выполнения операций сортировки. Однако, более эффективно такая процедура может выполняться в процессе сложения с помощью сортирующего устройства с применением логической фильтрации, идею которого предложил О.А. Ханджян для разрабатываемой им интервальной версии теории сигналов [7].

Функциональная схема такого логического сумматора показана на рис. 2, она представляет суммирующее и сортирующее устройство, сформированное на спаренных логических элементах И и ИЛИ. Светлые элементы на схеме соответствуют логическим элементам И, а зачерненные логическим элементам ИЛИ.

На вход суммирующего и сортирующего устройства одновременно поступают бинарные логические сигналы соответствующих уровней каждого из двух изображений. Размещение видеoinформации двух разных изображений по квантованным уровням вводится в произвольном порядке, что перемешивает её распределение по уровням и нарушает имевшуюся упорядоченную структуру уровней выходных сигналов входных изображений. Логическое устройство, показанное на рис. 2, перераспределяя размещение значений с 1 и 0 восстанавливает упорядоченную структуру их размещения по уровням для формирования суммарного изображения. Значения 1 при прохождении через такой фильтр концентрируются на нижних уровнях, а значения 0 соответственно на верхних.

Очевидно, что работа на уровне логических операций упрощает обработку сигналов. Нет необходимости учитывать переполнение и перенос разряда, следовательно, нет промежуточного запоминания и переноса,

то есть процедура обработки может выполняться в темпе поступления видеoinформации. Более того, скорость обработки данных даже увеличивается по сравнению с процессорной и остается постоянной. Вновь сформированный суммарный видеосигнал будет иметь число уровней квантования соответственно суммарному числу уровней совмещаемых изображений, что обеспечивает необходимое увеличение динамического диапазона.

### Технология логического увеличения динамического диапазона

Технология осуществления предлагаемого логического решения заключается в том, что при съемке получают не менее двух дискретизованных и квантованных на уровни изображений, которые формируют отображение наблюдаемой сцены в не менее двух разных, взаимно дополнительных частях требуемого динамического диапазона. Преобразуют представление значения амплитуд видеосигнала этих изображений из величин по уровням квантования в соответствующий этой величине набор единичных логических значений, размещенных по уровням квантования, заполняя остальные логическими нулями. Затем формируют общий, суммарный массив видеoinформации с суммарным для полученных снимков числом уровней квантования  $P$ . Для этого дополняют квантованные уровни одного изображения уровнями одного или более других изображений этой же сцены, полученных в дополняющих областях динамического диапазона в соответствии с координатами совмещаемых элементов изображения (пиксел). В полученном прямом совмещением общем видеoinформационном массиве уровни квантования отдельных составляющих изображений будут перемешаны произвольным образом. Поэтому для получения правильной структуры распределения видеoinформации по уровням квантования полученный массив сортируют. Операции сортировки производят по результатам выполнения логических операций И и ИЛИ между всеми парами уровней квантования. При этом результат операции ИЛИ между двумя уровнями ( $K$  и  $K+1$ ) относят к нижнему уровню ( $K$ ) в каждой паре, а результат операции И между теми же уровнями ( $K$  и  $K+1$ ) относят к верхнему уровню ( $K+1$ ). Этот цикл сравнений повторяют  $P-1$  раз, уменьшая число сравниваемых уровней при каждом шаге на 1. В результате сортировки получают видеoinформационный массив с общим числом уровней  $P$ , представленный упорядоченным размещением логических единиц по этим уровням. Для представления ее видеосигналом полученный набор логиче-

ских единиц для дискретизованных элементов снова преобразуют в амплитуду сигнала формируемого суммарного изображения с увеличенным динамическим диапазоном. При этом число уровней изображения будет соответствовать сумме числа уровне всех используемых входных изображений  $P$ . Описанное решение было защищено патентом РФ [8].

Выполнение предложенной процедуры совмещения видеоинформации в описанном способе не содержит собственно арифметической операции суммирования. Формирование суммарного изображения осуществляется в потоковом режиме сумматором на основе сортирующего устройства для уровней квантования общего массива полученной видеоинформации с применением логических функций. Все операции осуществляются применением каскада только логических операций И и ИЛИ.

### Экспериментальная проверка предложенного решения

Для иллюстрации особенностей решения поставленной задачи было выполнено её моделирование. Цифровой камерой последовательно с одной точки наблюдения получено два изображения одной и той же сцены при разных параметрах съемки. Как показано на рис. 3 а первое изображение отображает хорошо освещенные детали сцены, при этом детали в тени остались не проработанными. На рис. 3 б показано второе изображение, где детали в тени хорошо просматриваются, зато светлые участки сцены наоборот оказались не проработанными.

Многоступенчатая структура логического фильтра, показанная на рис. 2 была сформирована программным путем. В качестве входных сигналов использовались изображения, которые показаны на рис. 3 а и на рис. 3 б. Квантованные значения видеосигналов каждого из этих изображений были преобразованы в совокупность бинарных слоев видеоинформации и поступали на вход логического преобразователя в произвольном распределении уровней. На выходе блока – преобразователя было получено синтезированное изображение, представленное на рис. 3 в. Это цифровое изображение характеризуется увеличенным динамическим диапазоном и большим числом уровней, чем на исходных изображениях. Как хорошо видно на рис. 3 в, синтезированное изображение содержит всю информацию о деталях сцены, как первого, так и второго входных изображений.

Процедура многократных или параллельных наблюдений должна также уменьшать влияние шумов фотоприемника на получаемом изображении и, следовательно, повышать качество формируемого изображения. Однако, в данной работе уменьшение влияния шумов нами не оценивалось. Визуальное качество полученного в результате логического суммирования изображения, показанного на рис. 3 в, значительно повысилось по сравнению с показанными отдельными исходными изображениями.



а)



б)

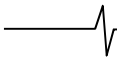


в)

Рис. 3. Изображения, полученные при разных условиях съемки (а, б) и результат их логического суммирования (в)

### Заключение

Многократное или параллельное наблюдение с параметрами, перекрывающими весь требуемый динамический диапазон отображаемой сцены, позволяет получить всю необходимую информацию о разрешаемых деталях изображения. Операцию синтеза суммарного изображения можно эффективно выполнять на основе использования только логических операций с применением описанного логического сумматора. Его структура легко реализуется аппаратно, например, на программируемых логических элементах и может функционировать в темпе поступления видеоинформации. Получаемое изображение будет содержать увеличенное число градаций соответственно суммарному числу градаций во взаимно дополняющих наблюдениях.



Предложенная схема решения относится к классу потоковых устройств обработки информации. Их отличительной особенностью является возможность выполнения преобразования сигналов в условиях непрерывного изменения наблюдаемой сцены. Технология параллельной обработки видеосигналов обеспечивает высокую эффективность предложенного способа. Применение для решения только логических операций сближает его с процедурами, которые выполняются в живых организмах. Как известно, в природных биологических организмах именно этот принцип позволяет создавать рекордные по быстродействию и сложности системы [9]. Описанное техническое решение может также служить примером возможного подхода к построению арифметических узлов компьютеров следующего поколения.

#### Литература

1. Бондаренко А.В., Докучаев И.В., Котцов В.А. Опыт работы цифровой фотокамеры на микроспутнике «Чибис-М» // Международная научно-техническая конференция «Академический микроспутник «Чибис-М». Результаты, уроки, перспективы». ИКИ РАН, 2014.

2. Гаранин С. Г., Зыков Л. И., Климов А. Н. и др. Дневное наблюдение звёзд слабой яркости (7 m-8 m) с равнинной местности. // Оптический журнал, т. 84, № 12, 2017.

3. «Союз-22» исследует Землю. / совместное издание АН СССР и Акад. наук ГДР – Москва; Berlin: Наука, 1980.

4. Виленчик Л.С., Гончаренко Б.Г., Курков И.Н и др. Способ расширения динамического диапазона передаваемых градаций яркости и/или освещенности в телевизионной системе. Патент РФ 2199827 // Бюллетень изобретений №6, 2003.

5. Карцев М.А. Арифметика цифровых машин. – М.: Наука, 1969.

6. Поспелов Д.А. Логические методы анализа и синтеза схем. – М: Энергия, 1968.

7. Ханджян О.А. Линейная фильтрация, основанная на теории симметрических функций // Радиотехника и электроника, вып. 8, 1986.

8. Котцов В.А. Способ получения изображений с увеличенным динамическим диапазоном. Патент РФ 2578799 // Бюллетень изобретений №9, 2016.

9. Бергсон А. Творческая эволюция. – М.: ТЕРРА-Книжный клуб, 2001.

УДК 621.396

## МОДИФИКАЦИЯ АРХИТЕКТУРЫ СВЕРТОЧНОЙ НЕЙРОННОЙ СЕТИ U-NET В ЗАДАЧЕ СЕГМЕНТАЦИИ МУЛЬТИКАНАЛЬНЫХ СПУТНИКОВЫХ ИЗОБРАЖЕНИЙ

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

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

*Павлов В.А., аспирант Ярославского государственного университета им. П.Г. Демидова, e-mail: vladimir@1pavlov.com;*

*Ларионов Р.В., магистрант Ярославского государственного университета им. П.Г. Демидова, e-mail: rv.larionov@yandex.ru.*

## MODIFICATION OF THE U-NET CONVOLUTIONAL NEURAL ARCHITECTURE IN THE MULTI-CHANNEL SATELLITE IMAGES SEGMENTATION PROBLEM

*Khryashchev V.V., Priorov A.L., Pavlov V.A., Larionov R.V.*

*The results of training a convolutional neural network for segmentation of four-channel satellite images are presented in this article. The algorithm is trained to segment the images of buildings and structures. We used a modification of the U-Net architecture convolutional neural network with two encoders. The values of segmentation quality metrics for 16 regions of the Russian Federation with different building densities are shown in the article. The training based on Spacenet images allowed us to achieve a Sorensen coefficient of 0,783 and a Jacquard index of 0,649.*

**Key words:** Earth remote sensing, segmentation, satellite images, convolutional neural network, deep learning.

**Ключевые слова:** дистанционное зондирование Земли, сегментация, спутниковые изображения, сверточная нейронная сеть, глубокое обучение.

### Введение

Задачи обнаружения и классификации различных объектов на двумерных данных, получаемых путем дистанционного зондирования земной поверхности, являются одновременно и классическими, и современными проблемами в области построения систем обработки визуальной информации [1-3]. Решение заключается в определении наличия или отсутствия конкретных признаков на спутниковых изображениях. После обнаружения этих признаков объект может быть дополнительно классифицирован как принадлежащий к одному из заранее определенных классов.

Большинство современных подходов к решению данной проблемы предполагают разработку алгоритмов глубокого обучения, а именно – использование сверточных нейронных сетей [4]. Например, построенный таким образом алгоритм автоматической сегментации позволяет быстро и достаточно точно выделять области интереса человека, например, сельскохозяйственные угодья, карьеры и другие объекты [5].

В настоящее время важной задачей является выделение зданий и сооружений на спутниковых изображениях. Интерес к указанной проблеме вызван увеличением пространственного разрешения двумерной информации, получаемой путем дистанционного зондирования земной поверхности, что делает возможным обнаружение сравнительно небольших объектов с линейными размерами порядка 5-10 метров.

*Представлены результаты обучения сверточной нейронной сети для сегментации четырехканальных спутниковых изображений. Алгоритм обучен выделять на изображениях здания и сооружения. Использовалась модификация архитектуры сверточной нейронной сети U-Net с двумя кодировщиками. Показаны значения метрик качества сегментации для 16 регионов Российской Федерации с различной плотностью застройки. Проведение предварительного обучения на базе изображений Spacenet позволило достичь значения коэффициента Серенсена 0,783 и индекса Жаккарда 0,649.*

ружение сравнительно небольших объектов с линейными размерами порядка 5-10 метров.

Анализ современной научно-технической литературы применительно к решению указанной задачи позволяет выявить особенности, предъявляемые к алгоритмам сегментации мультисканальных спутниковых изображений зданий и сооружений [6-10]:

– Объекты зданий и сооружения могут сильно различаться по своим линейным размерам: от небольших частных домов, размером в несколько пикселей, до огромных жилых домов в спальных районах. Разрабатываемый алгоритм должен одинаково хорошо выделять объекты разных линейных размеров. Данной цели можно достигнуть при использовании нескольких кодировщиков в составе сверточной нейронной сети.

– Сложным моментом является разделение объектов зданий и сооружений при их высокой плотности. Это достигается путем тщательного выбора функции потерь при обучении нейронной сети.

– Инвариантность к поворотам и цветовым искажениям, вызванным влиянием атмосферы при получении данных дистанционного зондирования земной поверхно-

сти. Эта проблема может быть решена путем аугментации (размножения) данных.

– Спутниковые изображения местности могут иметь различное пространственное разрешение. На практике существенным моментом является создание алгоритма, обладающего наилучшей обобщающей способностью.

– Спутниковые изображения могут быть очень объемными. Нередко размер файла превышает 150 МБ, а разрешение – больше 16000×16000 пикселей. Для упрощения реализации нейронной сети и увеличения скорости работы алгоритма изображения нередко разрезают на фреймы размером 512×512 пикселей и меньше. Итоговый алгоритм должен решать задачу обратной сборки полученных результатов.

– Алгоритм должен быть робастным к различного рода шумам и помехам. Причинами таких искажений могут быть: небольшие облака, блики от крыш зданий, изменение погодных условий, осадки и пр.

– На практике обычно получают мультиспектральные изображения в разных спектральных диапазонах: от ультрафиолетового до теплового инфракрасного. Использование одних и тех же методов для различных спектральных каналов не всегда приводит к наилучшему результату.

Автоматическая сегментация спутниковых изображений является важным этапом их обработки. В [11] показан алгоритм сегментации таких изображений с применением методов последовательного квадратичного программирования и анализа форм спектральных компонент. Сверточные нейронные сети позволяют получать признаки искомым объектам автоматически и классифицировать пиксели с учетом окружающей их области, то есть учитываются непосредственно свойства самого объекта [12]. Таким образом, повышается робастность алгоритма сегментации к шумам и помехам.

Современные нейронные сети позволяют детектировать в видеопотоке такие непростые объекты, как автомобили и пешеходы, практически в режиме реального времени (время детектирования составляет менее 100 мс) [13]. Структура модели сверточной нейронной сети параллельна, что делает идеальным применение

графических процессоров (GPU) для их обучения [14].

Использование сверточных нейронных сетей позволяет детектировать на мультиспектральных изображениях такие объекты как водные ресурсы, леса, облака над водой и над землей [15-17]. В работе [18] представлена архитектура нейронной сети U-Net, показавшая свою эффективность в решении задачи сегментации медицинских изображений. Сеть U-Net показала свою эффективность и в автоматической сегментации спутниковых изображений при малом количестве обучающих данных [19].

Целью данной работы является представление результатов обучения сверточной нейронной сети U-Net для сегментации спутниковых мультиспектральных изображений с целью выделения зданий и сооружений. Предложена модификация архитектуры данной нейронной сети. Работа является продолжением исследований авторов [20-21].

### Подготовка данных для обучения нейронной сети

Сверточная нейронная сеть предварительно обучена на базе изображений Spacenet [22], содержащей изображения 6 городских агломераций размером 650×650 пикселей: Рио-де-Жанейро (Бразилия), Лас-Вегас (США), Париж (Франция), Шанхай (Китай), Хартум (Судан) и Атланта (США). Восьмиканальные изображения получены спутниками WorldView-2 и WorldView-3 с различным пространственным разрешением. Spacenet содержит 2 подмножества общей площадью покрытия 3011 км<sup>2</sup> и 5555 км<sup>2</sup> для детектирования зданий и сооружений. Примеры изображений из базы Spacenet приведены на рис. 1.

Набор данных, на которых проходило дообучение модели, состоит из изображений 16 регионов Российской Федерации различной степени застройки. Каждое изображение имеет свою двоичную маску, размеченную вручную экспертами. Набор данных охватывает примерно 30 квадратных километров. Изображения являются четырехканальными: синий, зеленый, красный и ближний инфракрасный (NIR) каналы с пространственным разрешением 3 м/пиксель. Пример изображений в видимом диапазоне приведен на рис. 2.

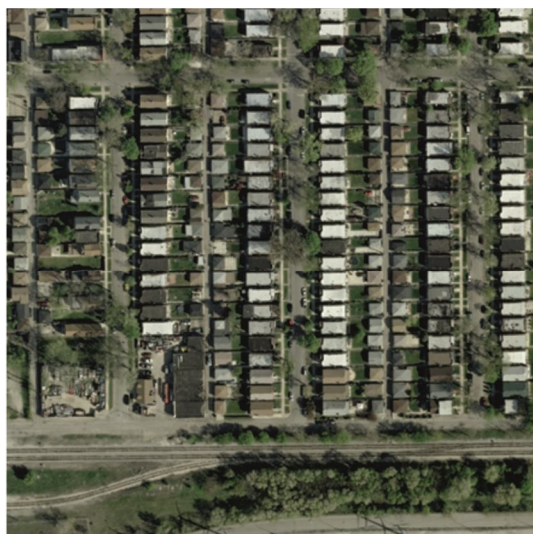


Рис. 1. Пример изображений базы Spacenet





Рис. 2. Пример изображений, использовавшихся для дообучения

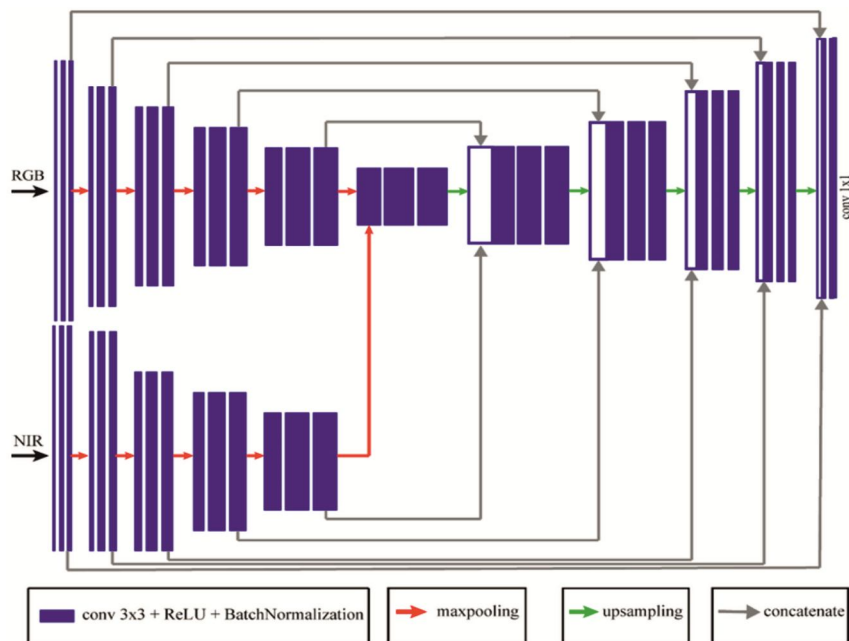


Рис. 3. Архитектура нейронной сети U-Net с двумя кодировщиками

Разработанная модель в качестве входных данных принимает фреймы размером  $256 \times 256$  пикселей. Для создания обучающей и тестовой выборки каждое изображение из базы данных разделялось на две непересекающиеся полосы (одна шириной 256 пикселей, вторая – все оставшиеся), затем каждая полоса разделялась на фрагменты  $256 \times 256$  пикселей с шагом 128, чтобы соседние фрагменты пересекались наполовину. Полоса шириной 256 пикселей предназначалась для тестов.

Для увеличения размера обучающей базы изображений проводилась аугментация данных. Она подразумевала выполнение следующих операций:

1. Повороты на 90, 180, 270 градусов и зеркальные отражения. Данные преобразования увеличили набор изображений в 8 раз.

2. Хроматические искажения. Изображения переводились из цветового пространства RGB в цветовое пространство HSV, и к новым цветовым координатам HSV добавлялись случайные значения. Для ближнего инфракрасного канала вместо хроматических искажений к нормализованному в интервале  $[0, 1]$  значениям добавлялись случайные величины из интервала  $[-0,06, +0,06]$ .

3. Случайные сдвиги, перемасштабирование и повороты на малые углы от  $-15$  до  $+15$  градусов.

В результате всех проведенных операций аугментации получена база из 9784 изображений, разрешением  $256 \times 256$  пикселей.

### Архитектура нейронной сети

В работе использовалась модификация известной архитектуры сверточной нейронной сети U-Net [18]. Оригинальная U-Net состоит из двух частей: кодировщика и декодировщика.

Кодировщик представляет собой сверточную нейронную сеть, состоящую из пяти блоков. Каждый такой блок состоит из двух сверточных слоев с фильтрами  $3 \times 3$ , с функцией активации ReLU и batch normalization, применяемой к каждому из них, а также слоя понижения дискретизации с окном  $2 \times 2$ . Декодировщик имеет такое же количество блоков, что и кодировщик. Каждый блок декодировщика состоит из слоя повышения дискретизации с окном  $2 \times 2$ , слияния с соответствующим набором признаков из кодировщика, двух сверточных слоев с фильтрами  $3 \times 3$  и функцией активации ReLU, примененной к

каждому из них. Последний слой использует сигмоидную функцию активации признаков для классификации на уровне пикселей [23].

Оригинальная архитектура сверточной нейронной сети U-Net была модифицирована: использовалось два кодировщика для RGB и NIR каналов с объединением их на центральном блоке (рис. 3). Итоговая нейронная сеть имела 47 сверточных слоев, 47 функций активации ReLU, 47 операций нормализации батчей, 1 функцию активации сигмоида, 10 операций повышения дискретизации, 5 операций понижения дискретизации, 11 операций слияния признаков.

### Численные результаты

Обучение нейронной сети проводилось на суперкомпьютере NVIDIA DGX-1, имеющемся в центре искусственного интеллекта и цифровой экономики ЯргУ им. П.Г. Демидова. Длительность процедуры обучения составила около 3 часов. В качестве оптимизатора для обучения модели использовался алгоритм Adam [24], в качестве функции потерь выбрана бинарная кросс-энтропия с оригинальными значениями весов и коэффициентом Серенсена.

Для оценки качества работы алгоритма сегментации использовались специальные метрики в задаче сегментации изображений: коэффициент подобия Серенсена (*dice*) и индекс Жаккарда (*IoU*).

Коэффициент подобия Серенсена принимает значения из отрезка [0, 1] и фактически является двоичной мерой сходства двух множеств. Данный показатель рассчитывается по следующей формуле:

$$dice = \frac{2I}{S},$$

где  $I = |X \cap Y|$  – мощность пересечения, а  $S = |X| + |Y|$  – сумма мощностей экспертной разметки  $X$  и предсказаний  $Y$ . В нашей задаче числитель  $I$  и знаменатель  $S$  можно вычислить по формулам:

$$I = \sum_{\substack{x \in X \\ y \in Y}} xy, \quad S = \sum_{\substack{x \in X \\ y \in Y}} (x + y),$$

где  $x, y \in [0, 1]$  – значения вероятностей принадлежности пикселей к тому или иному классу экспертной разметки  $X$  и предсказаний  $Y$  соответственно.

Индекс Жаккарда определяет степень сходства сравниваемых объектов и вычисляется по формуле:

$$IoU = \frac{|A \cap B|}{|A \cup B|} = \frac{|A \cap B|}{|A| + |B| + |A \cap B|},$$

где  $A$  и  $B$  – классы экспертной разметки и предсказаний нейронной сети, соответственно.

В задаче сегментации значение коэффициента Серенсена 0,7 является своего рода порогом. Превышение данного значения означает, что алгоритм хорошо справляется со своей задачей [2]. Исходя из этого, ставилась задача превысить данное значение метрики.

Вначале исследовалась целесообразность использования всех четырех каналов спутникового изображения. Для этого нейронная сеть независимо обучалась на RGB-изображениях, на NIR-изображениях и на 4-канальных изображениях, на которых изображена одна и та же территория, а именно 16 регионов РФ с различной степенью застройки. Использовалась классическая архитектура сети U-Net. Результаты представлены в табл. 1. Для удобства, использованы следующие сокращения типов местности: городская застройка – ГЗ, сельская местность – СМ, малый город – МГ.

Таблица 1. Значения коэффициента Серенсена и индекса Жаккарда в 16 регионах РФ для классической U-Net с RGB-, NIR- и 4-канальными изображениями на входе

Регион	Тип застройки	Коэффициент Серенсена			Индекс Жаккарда		
		RGB+NIR	RGB	NIR	RGB+NIR	RGB	NIR
1	ГЗ	<b>0,639</b>	0,586	0,580	<b>0,486</b>	0,436	0,419
2	ГЗ	<b>0,860</b>	0,713	0,773	<b>0,761</b>	0,591	0,641
3	СМ	<b>0,704</b>	0,617	0,547	<b>0,549</b>	0,457	0,392
4	ГЗ	<b>0,754</b>	0,670	0,707	<b>0,626</b>	0,535	0,567
5	МГ	<b>0,601</b>	0,572	0,429	<b>0,436</b>	0,407	0,275
6	МГ	<b>0,627</b>	0,592	0,551	<b>0,458</b>	0,421	0,381
7	СМ	0,582	<b>0,617</b>	0,550	0,443	<b>0,509</b>	0,418
8	СМ	0,669	0,715	<b>0,720</b>	0,633	0,672	<b>0,676</b>
9	СМ	<b>0,613</b>	0,520	0,612	0,466	0,38	<b>0,469</b>
10	СМ	<b>0,533</b>	0,479	0,493	<b>0,367</b>	0,321	0,320
11	МГ	<b>0,697</b>	0,651	0,674	<b>0,567</b>	0,509	0,540
12	ГЗ	<b>0,865</b>	0,847	0,821	<b>0,763</b>	0,735	0,697
13	ГЗ	<b>0,744</b>	0,689	0,708	<b>0,597</b>	0,530	0,553
14	ГЗ	<b>0,659</b>	0,643	0,519	<b>0,551</b>	0,531	0,395
15	ГЗ	0,536	0,527	<b>0,618</b>	0,400	0,395	<b>0,472</b>
16	ГЗ	0,588	<b>0,622</b>	0,619	0,456	<b>0,489</b>	0,478
Средние	–	<b>0,668</b>	0,633	0,627	<b>0,544</b>	0,506	0,478



Анализ результатов показывает, что использование 4-канальных изображений дает лучшие результаты в большинстве случаев. Более сложными наборами оказывается сельская местность (область маленьких зданий и сооружений), где использование всех 4-х каналов не дает преимущества. В то же время среднее значение коэффициента Серенсена остается ниже порога в 0,7. Для улучшения качества работы алгоритмов использовалось два варианта модификации классической сети U-Net. Первый из них заключается в использовании двух сетей U-Net для RGB- и NIR-каналов по отдельности с их объединением на последнем классифицирующем слое. Второй состоит в модификации U-Net с двумя кодировщиками, как показано на рис. 3. Сравнение работы данных алгоритмов по значениям коэффициента Серенсена изображено на рис. 4. Из приведенных данных следует, что схема с двумя кодировщиками является предпочтительной в большинстве случаев.

Дополнительно проведено предварительное обучение сети U-Net с двумя кодировщиками на базе изображения Spacenet. Во время предобучения значение коэффициента Серенсена достигло значения 0,84, индекса Жаккарда – 0,77.

Сравнение значений коэффициента Серенсена и индекса Жаккарда для классической архитектуры U-Net с 4-канальными изображениями на входе и модификации с двумя кодировщиками для RGB- и NIR-каналов без предобучения и с предобучением на базе изображений Spacenet приведены в табл. 2. Используются следующие сокращения: сеть с двумя кодировщиками – U-Net2, сеть с двумя кодировщиками с проведенным предварительным обучением на базе изображений Spacenet – U-Net2P. Как следует из приведенных данных – сеть U-Net2P показывает лучший результат в 13 случаях из 16, повышая при этом среднее значение коэффициента Серенсена до 0,783 и индекса Жаккарда до 0,649.

На рис. 5 и рис. 6 приведены примеры работы алгоритма в условиях городской застройки и сельской местности соответственно. Коэффициент Серенсена для полученного изображения составил 0,782 для городской застройки и 0,713 для изображения сельской местности. Меньшее значение коэффициента для регионов с сельской местностью связано с небольшим размером выделяемых объектов, которые зачастую занимают несколько пикселей при пространственном разрешении спутникового изображения – 3 метра на пиксель.

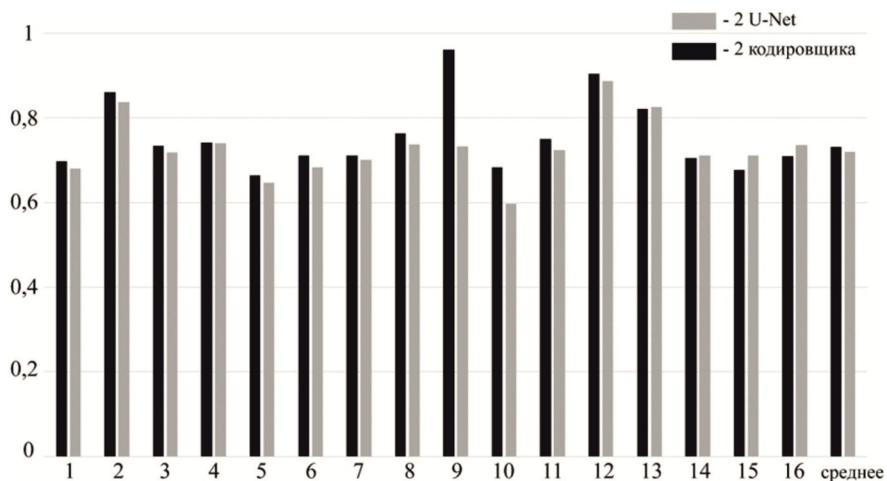


Рис. 4. Значения коэффициента Серенсена для сети с двумя кодировщиками и двумя U-Net при работе с 4-канальными изображениями



Рис. 5. Пример работы алгоритма в городской застройке. Коэффициент Серенсена для изображения равен 0,782

Таблица 2. Значения коэффициента Серенсена и индекса Жаккарда в 16 регионах РФ для классической U-Net, U-Net с двумя кодировщиками с предобучением и без предобучения

Регион	Тип застройки	Коэффициент Серенсена			Индекс Жаккарда		
		U-Net	U-Net2	U-Net2P	U-Net	U-Net2	U-Net2P
1	ГЗ	0,639	0,698	<b>0,768</b>	0,486	0,549	<b>0,638</b>
2	ГЗ	0,860	0,860	<b>0,887</b>	0,761	0,760	<b>0,803</b>
3	СМ	0,704	<b>0,733</b>	0,625	0,549	<b>0,583</b>	0,471
4	ГЗ	0,754	0,742	<b>0,842</b>	0,626	0,616	<b>0,738</b>
5	МГ	0,601	0,664	<b>0,667</b>	0,436	0,500	<b>0,504</b>
6	МГ	0,627	0,711	<b>0,737</b>	0,458	0,552	<b>0,584</b>
7	СМ	0,623	0,654	<b>0,747</b>	0,448	0,485	<b>0,597</b>
8	СМ	0,669	0,763	<b>0,797</b>	0,633	0,697	<b>0,745</b>
9	СМ	0,613	<b>0,753</b>	0,644	0,466	<b>0,637</b>	0,480
10	СМ	0,533	0,684	<b>0,714</b>	0,367	0,416	<b>0,582</b>
11	МГ	0,697	0,749	<b>0,880</b>	0,567	0,624	<b>0,788</b>
12	ГЗ	0,865	0,904	<b>0,916</b>	0,763	0,825	<b>0,846</b>
13	ГЗ	0,744	<b>0,821</b>	0,810	0,597	<b>0,700</b>	0,687
14	ГЗ	0,659	0,704	<b>0,770</b>	0,551	0,602	<b>0,672</b>
15	ГЗ	0,536	0,677	<b>0,716</b>	0,400	0,543	<b>0,587</b>
16	ГЗ	0,588	0,710	<b>0,783</b>	0,456	0,590	<b>0,671</b>
Средние	–	0,668	0,732	<b>0,783</b>	0,544	0,612	<b>0,649</b>



Рис. 6. Пример работы алгоритма в сельской местности. Коэффициент Серенсена для изображения равен 0,713

## Заключение

Описан процесс обучения модифицированной сверточной нейронной сети U-Net, предназначенной для сегментации мультиканальных спутниковых изображений. Обучение проводилось на базе из 9784 изображений разрешением 256×256 пикселей. Одна итерация обучения на суперкомпьютере NVIDIA DGX-1 составляла около 3 часов.

Показано, что использование ближнего инфракрасного канала повышает точность сегментации в задаче выделения зданий и сооружений. Предложено использование двух кодировщиков в сверточной нейронной сети при работе с 4-канальными изображениями.

Проведение предварительного обучения на базе изображений Spacenet позволило достичь значения

коэффициента Серенсена 0,783 и индекса Жаккарда 0,649.

Разработанный алгоритм сегментации может найти применение при решении задач по оценке уровня урбанизации различных регионов и отслеживанию строительства крупных объектов.

## Литература

1. Рассел Д. Спутниковая фотосъемка. М.: Мир, 2013. 157 с.
2. Визильтер Ю.В., Желтов С.Ю. Обработка и анализ изображений в задачах машинного зрения. Курс лекций и практических занятий. М.: Физматкнига, 2010. 672 с.
3. Алпатов Б.А., Ершов М.Д., Фельдман А.Б. Алгоритм обработки изображений для системы комбиниро-

ванного видения летательного аппарата // Цифровая обработка сигналов. 2015. № 3. С. 8-14.

4. Goodfellow Y., Bengio Y., Courville A. Deep Learning. The MIT Press, 2016, 800 p.

5. Барталев С.А., Ховратович Т.С. Анализ возможностей применения методов сегментации спутниковых изображений для выявления изменений в лесах // Современные проблемы дистанционного зондирования Земли из космоса. 2011. Т. 8. № 1. С. 44-62.

6. Zhang Z., Liu Q., Wang, Y. (2018). Road extraction by deep residual U-Net. IEEE Geoscience and Remote Sensing Letters, 15(5), pp. 749-753.

7. Zhang L., Zhang L., Du B. Deep learning for remote sensing data: A technical tutorial on the state of the art // IEEE Geoscience and Remote Sensing Magazine 4.2 (2016): pp. 22-40.

8. Chen Y., Jiang H., Li C., Jia X., Ghamisi P. Deep Feature Extraction and Classification of Hyperspectral Images Based on Convolutional Neural Networks, in IEEE Transactions on Geoscience and Remote Sensing, vol. 54, no. 10, pp. 6232–6251, Oct. 2016.

9. Zhu X., Tuia D., Mou L., Xia G-S., Zhang L., Xu F., Fraundorfer F. Deep Learning in Remote Sensing: A Comprehensive Review and List of Resources, in IEEE Geoscience and Remote Sensing Magazine, vol. 5, no. 4, pp. 8-36, Dec. 2017.

10. Iglovikov V., Mushinskiy S., & Osin V. (2017). Satellite imagery feature detection using deep convolutional neural network: A kaggle competition. arXiv preprint arXiv:1706.06169.

11. Григорьева О.В. Субпиксельная идентификация объектов местности по гиперспектральным данным на основе последовательного квадратичного программирования и метода анализа формы спектральных компонент // Цифровая обработка сигналов. 2018. № 3. С. 26-31.

12. Николенко С., Кадури А., Архангельская Е. Глубокое обучение. СПб.: Питер, 2019, 480 с.

13. Ерохин Д.Ю., Ершов М.Д. Современные сверточные нейронные сети для обнаружения и распознавания объектов // Цифровая обработка сигналов. 2018. № 3. С. 64-69.

14. Seferbekov S., Iglovikov V., Buslaev A., Shvets A. Feature Pyramid Network for Multi-Class Land Segmentation. Web: <https://arxiv.org/pdf/1806.03510.pdf>.

15. Иванов Е., Тищенко И., Виноградов А. Сегментация мультиспектральных снимков с применением сверточных нейронных сетей // Современные проблемы дистанционного зондирования Земли из космоса. 2019. Т. 16. № 1. С. 25-34.

16. Khryashchev V., Pavlov V., Priorov A., Ostrovskaya A. Deep Learning for Region Detection in High-Resolution Aerial Images // 2018 IEEE East-West Design & Test Symposium (EWDTS), Kazan, 2018, pp. 1-5.

17. Khryashchev V., Ivanovsky L., Pavlov V., Ostrovskaya A., Rubtsov A. Comparison of Different Convolutional Neural Network Architectures for Satellite Image Segmentation // Proceedings of the FRUCT'23, Bologna, Italy, 13-16 November 2018. pp. 172-179.

18. Ronneberger O., Fischer P., Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. Medical Image Computing and Computer-Assisted Intervention (MICCAI), Springer, LNCS, vol. 9351, 2015. pp. 234-241.

19. Соловьев Р., Тельпухов Д., Кустов А. Автоматическая сегментация спутниковых снимков на базе модифицированной сверточной нейронной сети UNET // Инженерный вестник Дона. 2017. Т. 47. № 4.

20. Ivanovsky L., Khryashchev V., Pavlov V., Ostrovskaya A. Building Detection on Aerial Images Using U-NET Neural Networks // 24th Conference of Open Innovations Association (FRUCT), Moscow, Russia, 2019, pp. 116-122.

21. Хрящев В.В., Приоров А.Л., Павлов В.А., Ивановский Л.И. Сегментация объектов на спутниковых изображениях с использованием сверточных нейронных сетей // Успехи современной радиоэлектроники. 2019. Т. 73. № 6. С. 28-34.

22. SpaceNet Database, Web: <http://explore.digitalglobe.com/spacenet>.

23. Gulli A., Pal S. Deep Learning with Keras, Packt Publishing, 2017, 320 p.

24. Kingma D.P., Ba J. Adam: A Method for Stochastic Optimization, Web: <https://arxiv.org/abs/1412.6980>.

## ИССЛЕДОВАНИЕ ПОДХОДОВ К ВЫДЕЛЕНИЮ КОНТУРОВ ОБЪЕКТОВ НА ИЗОБРАЖЕНИИ НА ОСНОВЕ ПРЕДВАРИТЕЛЬНОЙ ФИЛЬТРАЦИИ И НЕЧЕТКОЙ ЛОГИКИ

*Ершов М.Д., аспирант, ассистент кафедры автоматике и информационных технологий в управлении Рязанского государственного радиотехнического университета имени В.Ф. Уткина, e-mail: ershov.m.d@rsreu.ru; Георгиева С.С., студент Рязанского государственного радиотехнического университета имени В.Ф. Уткина, e-mail: frolova.sofia@yandex.ru.*

### RESEARCH ON APPROACHES TO OBJECT CONTOUR DETECTION ON THE BASIS OF PRELIMINARY FILTRATION AND FUZZY LOGIC

*Ershov M.D., Georgieva S.S.*

*The paper describes one of the problems in the field of image processing. This problem is associated with the detection of object edges on the observed scene. The aim of the work is to develop and study an algorithm for object contours extraction on the basis of pre-filtering and fuzzy logic. The proposed approaches were also compared with the well-known Sobel and Canny edge detectors. The use of pre-filtering is due to the ability to suppress the noise component as well as to emphasize the edges. The following approaches to preliminary filtering are considered: contrast enhancement, morphology-based image correction, bilateral and guided filtering. The developed contour detection algorithm is based on the gradient calculation and on the application of a fuzzy inference system. The fuzzy inference system is based on the Mamdani or Takagi-Sugeno model. The considered systems allow determining the degree of membership of a point to a contour or a homogeneous area. The results of experimental studies are shown. The studies consisted in processing the base of real images and reference contour images. The quality of the algorithm work was numerically estimated by calculating the precision, recall and F-measure.*

**Key words:** image processing, feature extraction, contours of objects, edge detection, pre-processing, filtering, fuzzy logic, Mamdani model, Takagi-Sugeno model.

**Ключевые слова:** обработка изображений, извлечение особенностей, контуры объектов, выделение границ, предварительная обработка, фильтрация, нечеткая логика, модель Мамдани, модель Такаги-Сугено.

#### Введение

Одной из фундаментальных проблем в области обработки изображений является выделение контуров объектов и других элементов наблюдаемой сцены. В общем случае выделение контуров применяется с целью значительного уменьшения количества данных на изображении, сохраняя при этом структурные свойства, которые могут быть использованы для дальнейшей обработки изображения [1]. Также контуры объектов и элементов сцены могут играть роль ключевых особенностей при связывании разнородных изображений, полученных, например, от датчиков разных типов [2].

Результатом выделения границ являются связанные кривые линии, которые и представляют контура объектов. На реальных изображениях далеко не всегда можно выделить все точки, принадлежащие контурам. Часто контурное изображение на выходе алгоритма имеет недостатки, например, отсутствие наблюдаемых границ, разрывы или наличие ложных границ, не принадлежащих объектам или сегментам на сцене. Задача выделения контуров значительно усложняется, если контур объекта размыт

*Рассмотрена одна из базовых задач в области обработки изображений, связанная с выделением границ объектов наблюдаемой сцены. Целью работы является исследование алгоритмов выделения контуров объектов на изображении на основе предварительной фильтрации и нечеткой логики, сравнение предлагаемых подходов с известными детекторами границ Собеля и Кэнни. Использование предварительной фильтрации обусловлено возможностью подавить шумовую составляющую изображения, а также подчеркнуть границы. В качестве подходов к предварительному улучшению изображения рассмотрены: повышение контраста, коррекция изображения с помощью морфологии, билатеральная и управляемая фильтрация. Разработанный алгоритм выделения контуров основан на вычислении градиента и на использовании системы нечеткого вывода. Система нечеткого вывода строится на основе модели Мамдани или Такаги-Сугено. Рассмотренные системы позволяют определить степень принадлежности точки к контуру или однородной области. Приведены результаты экспериментальных исследований, заключающихся в обработке базы натуральных изображений, имеющих эталонные контурные изображения. Качество работы алгоритма численно оценивалось путем расчета показателей точности, полноты и F-меры.*

или недостаточно гладкий, на что могут влиять шумы и различные особенности процесса регистрации изображения.

Задаче выделения контуров на изображении посвящено множество работ. Большинство алгоритмов осно-

вано на вычислении градиента (величины перепада яркости в определенном направлении) путем свертки изображения с ядром – градиентным оператором. Предложено большое число таких операторов: перекрестный оператор Робертса [3], оператор Собеля [4], оператор Превитт [5], дискретный оператор Лапласа [6], оператор Кирша [7], оператор Робинсона [8] и другие. Градиентный оператор может быть представлен одним или несколькими ядрами для вычисления градиента в разных направлениях, например, по вертикали, по горизонтали, в диагональных направлениях. К недостаткам применения того или иного градиентного оператора относятся высокая чувствительность к шумам и наличие разрывов в контурах.

Отдельно стоит отметить алгоритмы Марра-Хилдрет [9] и Кэнни [10], в которых предложен этап предварительной обработки входного изображения (размытие), а также сложные этапы пороговой обработки, на которых принимается окончательное решение о принадлежности точки контуру. Предварительная фильтрация уменьшает чувствительность к шуму, а адаптивная пороговая обработка позволяет устранить разрывы в контурах и уменьшить число ложных выделений. Также с целью выделения границ может применяться аппарат математической морфологии [11], подходы на основе генетического алгоритма [12] или преобразование Радона [13]. Последний подход в исходном виде предназначен для выделения прямых линий на изображении, но для преодоления этого ограничения предлагается разбивать изображение на небольшие блоки, в пределах которых кривизной линий можно пренебречь. Обнаруженные линии в каждом из блоков объединяются в длинные кривые или прямые линии с помощью многоагентного подхода. Подобные алгоритмы выделения контуров имеют преимущества в качестве выделения границ, но требуют дополнительных параметров и их точной настройки, а также могут быть весьма ресурсоемкими. Поэтому задача выделения контуров объектов является актуальной.

Целью данной работы является разработка алгоритма выделения контуров объектов на основе вычисления градиента и системы нечеткого вывода. В качестве системы нечеткого вывода может быть использована система Мамдани [14] или Такаги-Сугено [15]. Подход на основе нечеткой логики позволяет определить степень принадлежности точки к контуру или однородной области, путем пороговой обработки можно получить бинарное контурное изображение.

Также зачастую первым шагом алгоритма выделения контуров является предобработка изображения с целью улучшения конечного результата работы. К примеру, для подавления шума на изображении зачастую используется операция размытия. В данной работе будет проведено исследование качества работы алгоритма выделения контуров на исходных и предварительно обработанных изображениях. В качестве подходов к предварительному улучшению изображения рассмотрены: повышение контраста, коррекция изображения с помощью морфологии, фильтрация с подчеркиванием границ.

## Разработка алгоритма и программная реализация

В настоящей работе предлагается алгоритм выделения контуров, состоящий из двух этапов. Первым этапом алгоритма является предобработка изображения с целью улучшения конечного результата работы. Вторым этапом алгоритма выделения контуров является вычисление градиента и применение системы нечеткого вывода. Отметим, что предварительная обработка является необязательным этапом, а также может быть применена совместно с известными алгоритмами выделения границ.

Программная реализация и проведение экспериментальных исследований осуществлялись в среде разработки Matlab. Разработанная программа состоит из скрипта, в результате выполнения которого оценивается точность выделения контуров на множестве изображений при использовании заданного алгоритма предварительной фильтрации и разных алгоритмов выделения контуров. Главный скрипт использует пять вспомогательных функций:

1. Функция чтения изображения, которая осуществляет загрузку изображения и при необходимости преобразует его в градации серого.

2. Функция обработки изображения с выделением контуров, которая в зависимости от переданных аргументов применяет один из алгоритмов предварительной фильтрации, а затем – один из алгоритмов выделения контуров.

3. Функция для выделения контуров с помощью нечеткой логики, в которой реализовано создание и настройка заданной системы нечеткого вывода и ее применение.

4. Функция вывода результатов на экран, которая отвечает за отображение исходного изображения и полученного бинарного контурного изображения.

5. Функция расчета показателей качества выделения контуров на заданной выборке.

В созданной программе используются возможности Matlab, а именно функции модуля Image Processing Toolbox для чтения, отображения и обработки изображений. Также используется модуль Fuzzy Logic Toolbox, который позволяет проектировать и использовать системы нечеткого вывода двух типов: Мамдани и Такаги-Сугено. Используемые подходы к фильтрации изображений и выделению контуров будут рассмотрены в следующих разделах.

## Предварительная фильтрация изображения

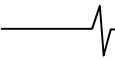
С целью повышения качества выделения границ рассматривались следующие способы предварительной фильтрации изображения:

1. Повышение контраста с помощью известных алгоритмов:

– повышение контраста с насыщением 1 % самых ярких и темных пикселей изображения (далее histAdjust);

– локальное повышение контраста с учетом границ (далее localContrast);

– глобальная эквализация гистограммы яркости изображения (далее histEq);



– адаптивная эквализация гистограммы яркости, выполняемая в разных блоках изображения (далее `adaptHistEq`).

2. Коррекция изображения с помощью морфологии (далее `morph`).

3. Использование специальных фильтров, подчеркивающих границы:

– билатеральный фильтр (далее `bilatFilter`);

– управляемая фильтрация изображения (далее `guidedFilter`).

Контраст можно понимать как разницу в яркостных характеристиках различных участков изображения. Повышение контраста обычно перераспределяет значения яркости в точках изображения в полный возможный диапазон яркостей. Изображение с хорошим контрастом имеет резкие различия между темными и светлыми участками. С целью коррекции контраста выбраны известные алгоритмы, которые выполняют или глобальную, или адаптивную локальную обработку изображения.

Алгоритм `histAdjust` является глобальным и состоит из следующих этапов:

1. Вычисление минимальной  $l_{in}$  и максимальной  $h_{in}$  яркости входного изображения. Для максимизации контраста строится гистограмма изображения, и отсекается заданный процент (в нашем случае 1 %) самых светлых и самых темных точек.

2. Вычисление яркости  $b'$  для каждой точки выходного изображения:

$$b' = (b - l_{in}) \cdot \frac{h_{out} - l_{out}}{h_{in} - l_{in}} + l_{out}, \quad (1)$$

где  $b$  – яркость точки входного изображения;  $l_{out}$  и  $h_{out}$  – заданные минимальная и максимальная яркости выходного изображения.

Алгоритм `localContrast` является локальным, увеличивает контрастность входного изображения, оставляя сильные края неизменными. Фильтрация изображений с учетом точек, относящихся к границам, может осуществляться на основе пирамиды Лапласа [16].

Еще одним подходом к улучшению контраста является эквализация – операция выравнивания гистограммы яркости изображения [17]. Для выполнения эквализации необходимо:

1. Построить гистограмму яркостей изображения  $h(b)$ .

2. Преобразовать гистограмму таким образом, чтобы все уровни яркости имели одинаковую частоту, а гистограмма соответствовала некоторой функции распределения.

Пусть  $h(b)$  – гистограмма яркостей  $b$ , тогда можно получить функцию распределения  $d(b)$ :

$$d(b) = \frac{1}{N} \sum_{i=0}^b h(i) = d(b-1) + \frac{1}{N} h(b), \quad (2)$$

где  $N$  – число точек изображения; значения яркости принадлежат диапазону [0; 255].

Подобные преобразования относятся к глобальным алгоритмам эквализации (`histEq`). Существуют локальные алгоритмы эквализации (`adaptHistEq`), которые зачастую дают лучшие результаты. Например, изображение можно разбить на блоки и выполнять выравнивание яркости в каждом из них, независимо друг от друга.

Одним из алгоритмов адаптивной эквализации является алгоритм CLAHE (`contrast-limited adaptive histogram equalization`) [18].

Алгоритм `morph` предполагает использование математической морфологии с целью предварительной обработки. Подход заключается в оценке фона с целью устранения неоднородности яркости, а также подчеркивания объектов. Выполняются следующие шаги:

1. К изображению  $I$  применяется операция морфологического открытия  $\circ$  со структурным элементом  $s$ , размер которого превышает размер объектов. В результате чего мы получаем подобие оценки фона.

2. Полученная оценка вычитается из исходного изображения.

3. Выполняется коррекция гистограммы для повышения контраста.

Первые два шага алгоритма также называются операцией «верх шляпы»:

$$I' = I - I \circ s. \quad (3)$$

Билатеральный фильтр (`bilatFilter`) – нелинейный фильтр, выполняющий пространственное усреднение в пределах своей маски [19]. Данный фильтр используется для подавления шума на изображении, но при этом сохраняет границы объектов. 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, например, яркости точек.

Управляемая фильтрация (`guidedFilter`) также выполняет сглаживание, сохраняя контуры на изображении [20]. При управляемой фильтрации обычно используется содержимое второго изображения, называемого направляющим изображением. В качестве направляющего мы будем использовать само исходное изображение, при этом фильтр учитывает особенности в соответствующей пространственной окрестности при вычислении значения яркости выходного пикселя. Т.к. управляющее изображение совпадает с фильтруемым, то границы объектов в результате будут совпадать с исходными.

## Выделение контуров на основе нечеткой логики

Основой разработанного алгоритма выделения контуров является использование системы нечеткого вывода. В качестве системы нечеткого вывода может быть использована система Мамдани или Такаги-Сугено. Разработанный подход сравнивался с известными детекторами границ Кэнни и Собеля. В дальнейшем будут использованы следующие обозначения:

1. `canny` – детектор границ Кэнни.

2. `sobel` – детектор границ на основе градиентного оператора Собеля.

3. `edgeMamdani` – выделение контуров на основе системы нечеткого вывода Мамдани.

4. `edgeTG` – выделение контуров на основе системы нечеткого вывода Такаги-Сугено.

Алгоритм работы в случае использования систем нечеткого вывода состоит из следующих этапов:

1. Вычисление градиента изображения.

– в случае Мамдани вычисляются градиенты изоб-

ражения по горизонтали и вертикали путем свертки изображения с градиентными фильтрами  $G_x$  и  $G_y$ ;

– для системы Такаги-Сугено вычисляются модули градиента.

2. Оценка выхода детектора границ для каждой строки пикселей изображения путем подачи строки на вход системы нечеткого вывода.

3. Выбор порога и дальнейшая пороговая обработка.

– для системы Мамдани порог устанавливался как минимальное значение в выходном массиве, умноженное на 2;

– для Такаги-Сугено был установлен фиксированный порог, равный 0,1.

Алгоритм нечеткого вывода Такаги-Сугено отличается от алгоритма Мамдани способом формирования заключений из правил нечетких продукций и формулой для получения четкого значения выхода нечеткой модели. В системах типа Мамдани база знаний состоит из правил вида:

Если  $a = 0$  и  $b = 0$ , то  $c =$  «белый». (4)

В системах типа Такаги-Сугено база знаний состоит из правил вида:

Если  $a = 0$  и  $b = 0$ , то  $c = k_0 + k_1 \cdot a + k_2 \cdot b$ . (5)

Таким образом, различие заключается в разных способах задания значений выходной переменной в правилах, образующих базу знаний. В системах типа Мамдани значения выходной переменной задаются нечеткими терминами, в системах типа Такаги-Сугено – как линейная комбинация входных переменных.

Этап определения системы нечеткого вывода для выделения контуров состоит из следующих шагов:

1. Создание объекта-системы нечеткого вывода на основе одной из моделей.

2. Указание градиентов изображения в качестве входных переменных системы.

3. Задание функции принадлежности для каждого входа (градиенты  $I_x$  и  $I_y$ ). Используются гауссовские функции принадлежности с нулевым средним и СКО 0,05.

4. Указание выходной переменной системы (вектор-строка, т.к. обработка выполняется построчно).

5. Задание функции принадлежности для выхода. Для системы Мамдани используется треугольная функция принадлежности, а для Такаги-Сугено – линейная, что связано с описанными особенностями работы данной нечеткой модели.

6. Задание правил для системы нечеткого вывода. Пиксель будет белым (*white*), если он принадлежит однородной области, и черным (*black*) – в противном случае. Пиксель находится в однородной области, когда градиент изображения равен нулю в обоих направлениях. Если в любом направлении имеется ненулевой градиент, то определяется степень принадлежности пикселя контуру.

Для проведения экспериментальных исследований была проведена настройка функций принадлежности для входов и выходов систем нечеткого вывода. Так как в результате исследований требовалось добиться как можно больших значений показателей качества, то проводилась эмпирическая настройка параметров функций принадлежности. Однако отклонение данных параметров от исходных значений не привело к существенному увеличению качества работы алгоритма, а при значительных отклонениях качество наоборот ухудшалось.

Используемые функции принадлежности для выхода ( $I_{out}$ ) системы Мамдани приведены на рис. 1.

Выход системы Такаги-Сугено определяется согласно следующим выражениям:

$$white = 0 \cdot I_x + 0 \cdot I_y + 1, \quad (6)$$

$$black = -0,5 \cdot I_x - 0,5 \cdot I_y + 0. \quad (7)$$

#### Экспериментальные исследования

При проведении экспериментальных исследований использовались изображения из базы Berkeley Segmentation Data Set and Benchmarks 500 (BSDS500) [21]. Данная база содержит большой набор натуральных изображений, которые были сегментированы людьми. Результаты ручной разметки берутся в качестве эталонных данных

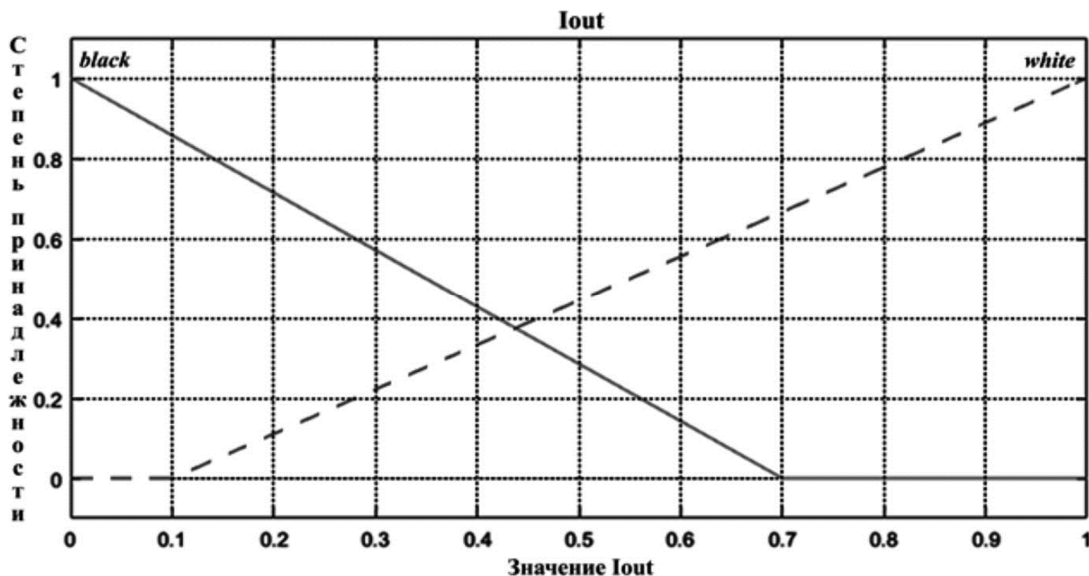


Рис. 1. Функции принадлежности для выхода системы Мамдани



при сравнении различных алгоритмов обнаружения контуров. Для тестирования было выбрано 30 изображений, отличающихся фоновой составляющей, числом и размерами объектов.

Численно оценить качество работы алгоритма выделения контуров можно путем расчета показателей точности (*precision*) и полноты (*recall*), а также *F*-меры. Точность и полнота являются метриками, которые часто используются при оценке большого числа алгоритмов извлечения информации и классификации. Они могут быть использованы сами по себе, а также в качестве базиса для производных метрик, например, *F*-меры.

Точность, полнота и *F*-мера рассчитываются по формулам:

$$precision = \frac{n_{TP}}{n_{TP} + n_{FP}}, \quad (8)$$

$$recall = \frac{n_{TP}}{n_{TP} + n_{FN}}, \quad (9)$$

$$F = 2 \frac{precision \cdot recall}{precision + recall}, \quad (10)$$

где  $n_{TP}$  (True Positive) – число точек, верно отнесенных к границам объектов;  $n_{FP}$  (False Positive) – число точек,

ложно отнесенных к границам;  $n_{FN}$  (False Negative) – число необнаруженных точек границ.

В табл. 1 приведено сравнение показателей качества, полученных для известных алгоритмов (*canny*, *sobel*) и для разработанного алгоритма (*edgeMamdani*, *edgeTG*). Жирным шрифтом выделены наилучшие результаты для каждого алгоритма. Для разработанного алгоритма использовались простые градиентные фильтры  $G_x$  и  $G_y$ :

$$G_x = [-1 \ 1] \text{ и } G_y = G_x^T. \quad (11)$$

Расшифровка обозначений, используемых в таблице, дана в разделах «Предварительная фильтрация изображения» и «Выделение контуров на основе нечеткой логики».

Полученные результаты содержат низкие значения показателя точности, что обусловлено выделением большого числа точек, не принадлежащих контурам объектов. Для уменьшения чувствительности разработанного алгоритма к шуму был добавлен шаг сглаживания. Изображение перед обработкой подвергалось размытию фильтром Гаусса с размером 3×3 пикселя. Также влияние на работу алгоритма оказал вид градиентных фильтров.

Таблица 1 – Сравнение № 1 показателей точности, полноты и *F*-меры

Алгоритм	Точность	Полнота	<i>F</i> -мера
<i>canny</i>	0,310465	0,969709	0,470343
histAdjust и <i>canny</i>	0,309696	0,970775	0,469585
localContrast и <i>canny</i>	0,288703	0,978149	0,445821
histEq и <i>canny</i>	0,276131	0,958293	0,428726
adaptHistEq и <i>canny</i>	0,268695	0,982028	0,421942
morph и <i>canny</i>	0,289668	0,944949	0,443411
bilatFilter и <i>canny</i>	0,335719	0,964572	0,498081
<b>guidedFilter и <i>canny</i></b>	<b>0,399909</b>	<b>0,882148</b>	<b>0,550332</b>
<i>sobel</i>	0,459714	0,677770	0,547841
histAdjust и <i>sobel</i>	0,459105	0,679559	0,547991
localContrast и <i>sobel</i>	0,451945	0,693517	0,547257
histEq и <i>sobel</i>	0,453389	0,710972	0,553688
adaptHistEq и <i>sobel</i>	0,425707	0,703043	0,530304
morph и <i>sobel</i>	0,467227	0,638380	0,539556
<b>bilatFilter и <i>sobel</i></b>	<b>0,498501</b>	<b>0,626671</b>	<b>0,555285</b>
guidedFilter и <i>sobel</i>	0,501015	0,594953	0,543958
<i>edgeMamdani</i>	0,233804	0,944485	0,374821
histAdjust и <i>edgeMamdani</i>	0,212279	0,964543	0,347975
localContrast и <i>edgeMamdani</i>	0,233804	0,944485	0,374821
histEq и <i>edgeMamdani</i>	0,183596	0,967498	0,308625
adaptHistEq и <i>edgeMamdani</i>	0,171201	0,981448	0,291545
morph и <i>edgeMamdani</i>	0,235807	0,937294	0,376814
bilatFilter и <i>edgeMamdani</i>	0,338832	0,853451	0,485080
guidedFilter и <i>edgeMamdani</i>	0,358880	0,776035	0,490792
<i>edgeTG</i>	0,359921	0,744742	0,485303
histAdjust и <i>edgeTG</i>	0,330639	0,848946	0,475921
localContrast и <i>edgeTG</i>	0,359921	0,744742	0,485303
histEq и <i>edgeTG</i>	0,243753	0,882773	0,382021
adaptHistEq и <i>edgeTG</i>	0,207522	0,939614	0,339960
morph и <i>edgeTG</i>	0,363248	0,694546	0,477016
bilatFilter и <i>edgeTG</i>	0,432715	0,627801	0,512314
guidedFilter и <i>edgeTG</i>	0,444555	0,498321	0,469905



Таблица 2 – Сравнение № 2 показателей точности, полноты и F-меры

Алгоритм	Точность	Полнота	F-мера
edgeMamdani	0,458662	0,748729	0,568852
histAdjust и edgeMamdani	0,428691	0,847354	0,569342
localContrast и edgeMamdani	0,458662	0,748729	0,568852
histEq и edgeMamdani	0,333163	0,876547	0,482815
adaptHistEq и edgeMamdani	0,288818	0,948700	0,442824
morph и edgeMamdani	0,465672	0,681301	0,553218
<b>bilatFilter и edgeMamdani</b>	<b>0,502195</b>	<b>0,704731</b>	<b>0,586469</b>
guidedFilter и edgeMamdani	0,534150	0,611038	0,570013
edgeTG	0,412318	0,715640	0,523195
histAdjust и edgeTG	0,384198	0,824924	0,524238
localContrast и edgeTG	0,412318	0,715640	0,523195
histEq и edgeTG	0,291865	0,870162	0,437115
adaptHistEq и edgeTG	0,244143	0,938963	0,387524
morph и edgeTG	0,434588	0,646063	0,519633
<b>bilatFilter и edgeTG</b>	<b>0,448021</b>	<b>0,638523</b>	<b>0,526572</b>
guidedFilter и edgeTG	0,461932	0,523902	0,490969

Таблица 3 – Сравнение № 3 показателей точности, полноты и F-меры

Алгоритм	Точность	Полнота	F-мера
edgeMamdani	0,496534	0,715352	0,586188
<b>histAdjust и edgeMamdani</b>	<b>0,467771</b>	<b>0,817761</b>	<b>0,595123</b>
localContrast и edgeMamdani	0,496534	0,715352	0,586188
histEq и edgeMamdani	0,370054	0,860604	0,517561
adaptHistEq и edgeMamdani	0,315028	0,937452	0,471582
morph и edgeMamdani	0,506396	0,640841	0,565741
<b>bilatFilter и edgeMamdani</b>	<b>0,522625</b>	<b>0,683224</b>	<b>0,592230</b>
guidedFilter и edgeMamdani	0,547305	0,593177	0,569318
edgeTG	0,457267	0,643262	0,534548
<b>histAdjust и edgeTG</b>	<b>0,440386</b>	<b>0,763974</b>	<b>0,558709</b>
localContrast и edgeTG	0,457267	0,643262	0,534548
histEq и edgeTG	0,352177	0,831134	0,494724
adaptHistEq и edgeTG	0,294726	0,913882	0,445711
morph и edgeTG	0,493093	0,563239	0,525837
bilatFilter и edgeTG	0,469536	0,581134	0,519408
guidedFilter и edgeTG	0,469290	0,473844	0,471556

При использовании системы Мамдани с градиентами, рассчитанными на расширенной области, толщина контуров увеличивается, поэтому полученное бинарное контурное изображение подвергалось операции эрозии. Проводились исследования доработанного алгоритма с разными градиентными фильтрами. В табл. 2 представлены показатели качества, полученные при использовании фильтров  $G_x$  и  $G_y$ :

$$G_x = [-0,5 \quad -0,5 \quad 0,5 \quad 0,5] \text{ и } G_y = G_x^T. \quad (12)$$

В табл. 3 представлены показатели качества, полученные при использовании фильтров  $G_x$  и  $G_y$ :

$$G_x = \frac{1}{6} \cdot \begin{bmatrix} -1 & -1 & 1 & 1 \\ -1 & -1 & 1 & 1 \\ -1 & -1 & 1 & 1 \end{bmatrix} \text{ и } G_y = G_x^T. \quad (13)$$

Примеры результата выделения контуров объектов сравниваемыми алгоритмами приведены на рис. 2. С целью улучшения визуализации контура были утолщены и обозначены черными пикселями.

Из анализа полученных результатов можно сделать следующие выводы. Использование предварительной обработки на основе повышения контраста перераспределяет значения яркости в точках изображения в полный возможный диапазон яркостей. Изображение с хо-

рошим контрастом должно иметь резкие различия между темными и светлыми участками. Однако, исходя из результатов сравнения показателей качества, можно сделать вывод, что алгоритм выделения контуров с использованием предварительной обработки с повышением контраста зачастую работает хуже, чем без его применения. Только при увеличении области, захватываемой градиентными фильтрами, удалось добиться некоторого повышения показателей качества (1,5-4,5 % по F-мере).

Использование математической морфологии для предварительной обработки позволяет оценить фоновую составляющую и затем, например, устранить неоднородность освещения, выделить фрагменты, принадлежащие объектам. При выделении контуров с использованием предварительной обработки изображения на основе морфологии иногда удается добиться незначительного повышения показателя точности, но показатель полноты при этом заметно уменьшается (на 3-14 %), следовательно, значение F-меры также снижалось (на 2-6 %). При использовании морфологии результат может зависеть от размера структурного элемента (в данном случае 30 пикселей – радиус дискообразного структурного элемента). Выбирать размер структурного элемента следует исходя из размера объектов на изображении, но в общем случае мы заранее

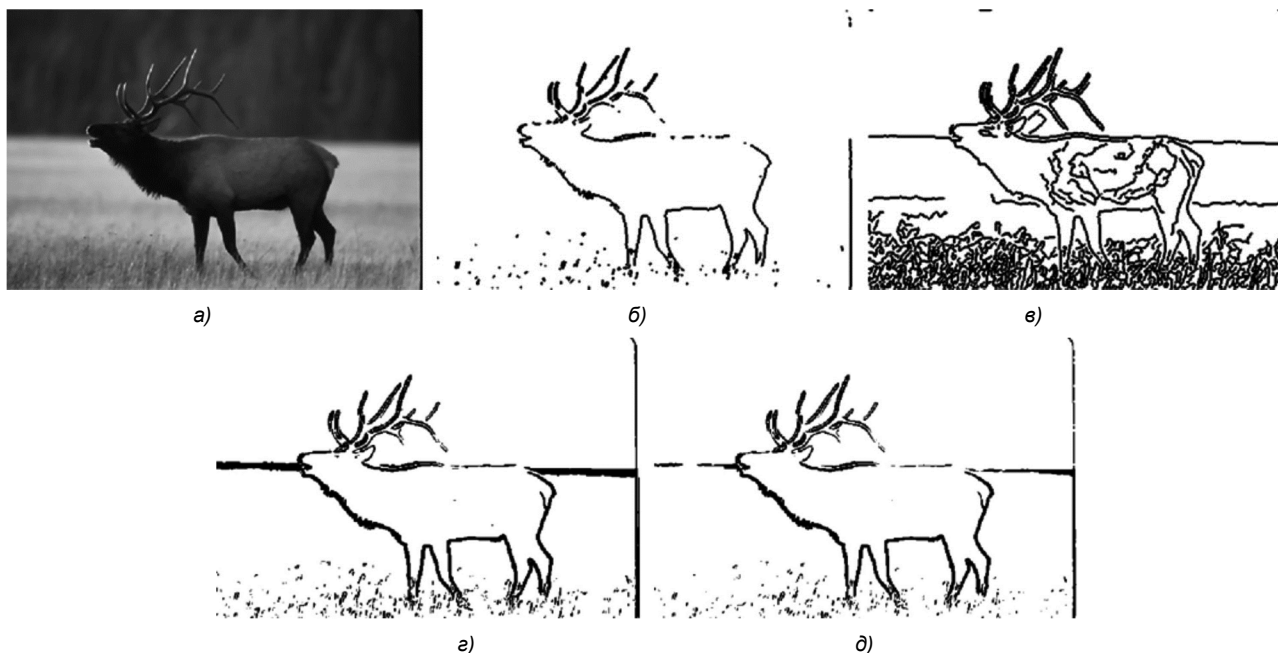


Рис. 2. Результаты выделения контуров:

а) исходное изображение; б) оператор Собеля; в) алгоритм Кэнни;  
 г) алгоритм на основе системы Мамдани; д) алгоритм на основе системы Такаги-Сугено

не знаем размеры объектов, а также на одной сцене могут быть представлены объекты, значительно отличающиеся по размеру.

При использовании билатерального и управляемого фильтров для предварительной обработки удалось добиться повышения показателей качества в большинстве случаев. Данный подход показал себя лучше всего, при этом увеличение точности (на 5-23 %) с одновременным уменьшением полноты (на 5-14 %) приводило к достижению баланса между этими метриками, а итоговое значение  $F$ -меры оказывалось выше (на 6-15 %), чем без предварительной обработки.

При выделении контуров на изображении с использованием систем нечеткого вывода удалось получить более высокие значения показателей качества по сравнению с операторами Собеля (на 7-11 % по  $F$ -мере) и Кэнни (на 8-21 % по  $F$ -мере). При этом для борьбы с проблемой выделения большого числа точек, не принадлежащих контурам объектов (и, следовательно, низким показателем точности), в итоге применялся сглаживающий фильтр. Также на работу подходов на основе как системы Мамдани, так и Такаги-Сугено значительное влияние оказывает вид градиентных фильтров, с увеличением размера фильтра удалось добиться и увеличения показателей качества.

### Заключение

В статье предложены подходы к выделению контуров объектов с применением предварительной фильтрации и систем нечеткой логики Мамдани или Такаги-Сугено. При проведении экспериментальных исследований численно оценено качество работы алгоритмов путем расчета показателей точности, полноты и  $F$ -меры. Предложенные подходы были сравнены с известными детекторами границ Собеля и Кэнни.

При сравнении результатов выделения контуров с эталонами, размеченными вручную, показатель точности оказался достаточно низок, так как базовые алгоритмы выделяют много лишних точек, принадлежащих к незначительным объектам сцены, или возникающих из-за шума. При использовании предложенных подходов удалось добиться увеличения показателя точности. В качестве алгоритма предварительной обработки лучше всего показали себя алгоритм управляемой фильтрации и билатеральный фильтр, в отдельных случаях коррекция гистограммы яркости также приводила к высоким показателям.

Из анализа результатов экспериментов следует, что алгоритм выделения контуров на основе системы нечеткого вывода Мамдани может давать более высокие значения показателей качества по сравнению с операторами Собеля и Кэнни. Хотя алгоритм на основе системы Мамдани и получил более высокие показатели качества, следует отметить, что алгоритм на основе системы Такаги-Сугено работает в 4-5 раз быстрее. Предложенные подходы могут быть использованы в системах технического зрения при решении таких прикладных задач, как пилотирование автомобильного и воздушного транспорта, для повышения скорости и точности методов дактилоскопии, обнаружения объектов для слежения и видеоаналитики, распознавания чертежей, текстов, таблиц и т.п.

Исследования выполнены при финансовой поддержке гранта Президента Российской Федерации МК-2737.2019.9 (соглашение 075-15-2019-350).

### Литература

1. 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.

2. Алпатов Б.А., Ершов М.Д., Фельдман А.Б. Алгоритм обработки изображений для системы комбинированного видения летательного аппарата // Журнал «Цифровая обработка сигналов», 2015. № 3. С. 8-14.
3. Roberts L.G. Machine Perception of 3-D Solids. MIT Press, 1965. 82 p.
4. Sobel I., Feldman G. A 3x3 Isotropic Gradient Operator for Image Processing // Stanford Artificial Intelligence Project. 1968.
5. Prewitt J.M.S. Object Enhancement and Extraction, Picture processing and Psychopictorics. New York: Academic Press, 1970. P. 75-149.
6. Титов И.О., Емельянов Г.М. Выделение контуров изображения движущегося объекта // Вестник Новгородского государственного университета, 2010. № 55. С. 27-31.
7. Kirsch R.A. Computer determination of the constituent structure of biological images // Computers and Biomedical Research, 1971. Vol. 4. P. 315-328.
8. Robinson G.S. Edge detection by compass gradient mask // Computer Graphics and Image Processing, 1977. Vol. 6. P. 492-501.
9. Marr D.C., Hildreth E.C. Theory of Edge Detection // Proceedings of the Royal Society of London. Series B, Biological Sciences, 1980. Vol. 207. P. 187-217.
10. Canny J.F. A Computational Approach to Edge Detection // IEEE Transactions on Pattern Analysis and Machine Intelligence, 1986. Vol. 8. P. 679-698.
11. Визильтер Ю.В., Желтов С.Ю. Обработка и анализ изображений в задачах машинного зрения. Курс лекций и практических занятий. М.: Физматкнига, 2010. 672 с.
12. Mitchell M. An introduction to genetic algorithm. MIT Press, 1996. 158 p.
13. Babayan P.V., Shubin N.Y. Line detection in a noisy environment with weighted Radon transform // Proc. SPIE 9024, Image Processing: Machine Vision Applications VII, 2014. Vol. 902409. 6 p.
14. Mamdani E.H., Assilian S. An experiment in linguistic synthesis with a fuzzy logic controller // International Journal of Man-Machine Studies, 1975. Vol. 7. No. 1. 13 p.
15. Sugeno M. Industrial applications of fuzzy control. Elsevier Science Ltd, 1985. 278 p.
16. Paris S., Hasinoff S.W., Kautz J. Local Laplacian Filters: Edge-Aware Image Processing with a Laplacian Pyramid // Communications of the ACM, 2015. Vol. 58. No. 3. P. 81-91.
17. Szeliski R. Computer Vision: Algorithms and Applications. Springer, 2010. 979 p.
18. Zuiderveld K. Contrast Limited Adaptive Histogram Equalization // Graphic gems IV, 1994. P. 474-485.
19. Tomasi C., Manduchi R. Bilateral filtering for gray and color images // 6th International conference on computer vision. IEEE, 1998. P. 839-846.
20. He K., Sun J., Tang X. Guided Image Filtering // IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013. Vol. 35. P. 1397-1409.
21. Arbelaez P., Maire M., Fowlkes C., Malik J. Contour Detection and Hierarchical Image Segmentation // IEEE TPAMI, 2011. Vol. 33. No. 5. P. 898-916.

## УЛУЧШЕНИЕ КАЧЕСТВА ТРАДИЦИОННОГО МЕТОДА КОМПЕНСАЦИИ ДВИЖЕНИЯ

Дам Чонг Нам, аспирант Московского физико-технического института (национального исследовательского университета), e-mail: chong.dam@phystech.edu.

## THE QUALITY IMPROVEMENT OF THE TRADITIONAL MOTION COMPENSATION METHOD

*Dam Trong Nam*

*This paper is dedicated to the research and development of prediction quality improvement of traditional motion compensation method for video codecs based on inter-frame block motion compensation. The paper proposes a new motion compensation method with the use of approximation functions with additional parameters. For applying the considered approximation functions in the video codec, the subtasks associated with the accuracy and the transmission method of additional parameters have been solved. Based on the results of high-definition video processing with the approximation functions studied, the best suitable function has been chosen. Application of the proposed method significantly reduces the amount of compressed data (from 15 % to 34 %) at a given quality of the reconstructed frame for high-definition videos.*

**Key words:** video coding, approximation function, motion estimation, motion compensation.

**Ключевые слова:** видеокodирование, функция аппроксимации, анализ движения, компенсация движения.

### Введение

Для уменьшения временной избыточности в видеокodеках используется анализ и компенсация движения. Традиционным методом компенсации движения является метод сопоставления блоков (СБ), заключающийся в нахождении блоков опорного кадра, соответствующих блокам текущего кадра с использованием какой-либо метрики [1]. В современных реализациях видеокodеков х.264 [2] и х.265 [3], а также в вейвлет-видеокodеке Dirac [4] использована простая метрика суммы абсолютных разностей (sum of absolute differences, SAD) из-за легкости вычисления, что не гарантирует оптимальность анализа движения. Действительно, традиционный метод СБ с метрикой SAD удачно работает для несложных видео, когда плоско-параллельное перемещение (трансляция) является основным типом движения объектов текущего кадра относительно опорного кадра, а для более сложных видео, в которых происходят изменение масштаба (зуммирование), вращение, изменение яркости и т.п., такой традиционный метод оказывается неэффективным.

Пусть  $Y$  – значение яркости какой-то точки рассматриваемого блока размером  $M \times N$  в текущем кадре,  $X$  – значение яркости точки с таким же положением в блоке соответствия в опорном кадре, тогда традиционный метод компенсации движения использует простую функцию аппроксимации  $Y = X$  для предсказания рассматриваемого блока текущего кадра на основе опорного кадра и вектора движения, который определяется условием минимизации суммы модулей разности для всех точек блока:

*Работа посвящена исследованию и разработке способа повышения качества предсказания в традиционном методе компенсации движения для видеокodеков, основанных на межкадровой блочной компенсации движения. В работе предложен новый метод компенсации движения с использованием функции аппроксимации с дополнительными параметрами. Для применения исследованных функций аппроксимации в видеокodеке решены подзадачи, связанные с точностью и способом передачи дополнительных параметров. На основе результатов обработки видео высокого разрешения с исследованными функциями аппроксимации выбрана наилучшая функция, подходящая для поставленной цели. Применение предложенного метода позволяет существенно снизить объем компрессированных данных (от 15 % до 34 %) при заданном качестве восстановленного кадра для видео высокого разрешения.*

$$SAD(V_x, V_y) = \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} |Y(x+m, y+n) - X(x+V_x+m, y+V_y+n)|. \quad (1)$$

Нужно отметить, что современные технологии и возможность распараллеливания вычислений позволяют выполнить процесс анализа движения с использованием более сложных, но более точных функций аппроксимации. В данной работе предложен способ повышения качества традиционного метода компенсации движения, заключающийся в применении других функций аппроксимации.

Целью работы является исследование различных функций аппроксимации для задачи сжатия видео, разработка улучшенного варианта традиционного метода компенсации движения и его применение для вейвлет-видеокodека.

Этот новый подход ранее не применялся. В стандарте H.265 [5] для частного случая, когда яркость кадра глобально уменьшается или увеличивается, применяется функция аппроксимации  $Y = aX$ , однако параметр  $a$

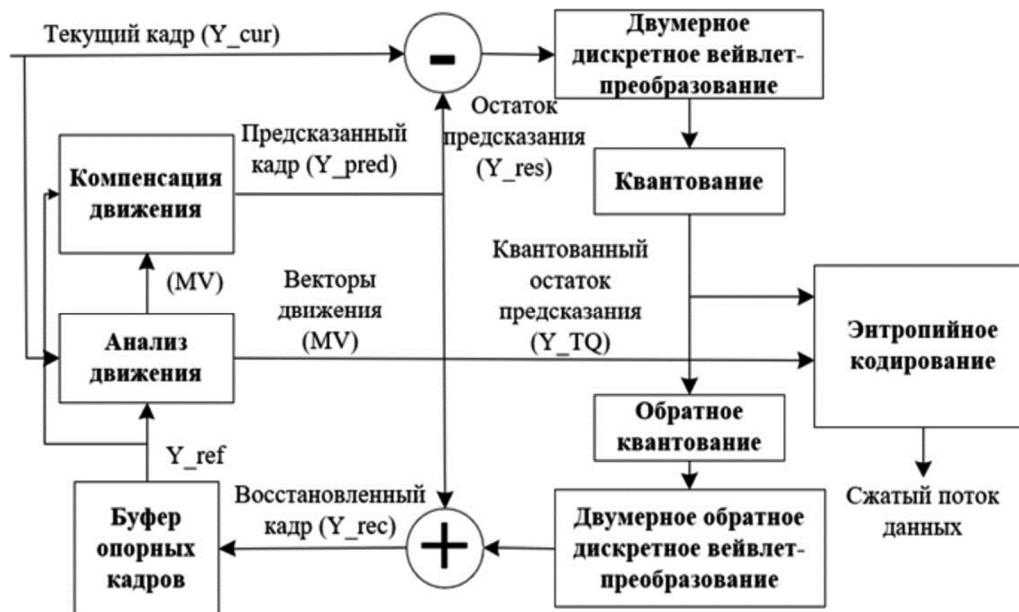


Рис. 1. Общая схема вейвлет-видеокодера

принимает одно значение для всего кадра, а не различные значения для каждого блока, как в предлагаемом ниже методе.

Практическая значимость работы подтверждена результатами обработки видео высокого разрешения 1080 p (1920×1080). Применение предложенного метода позволит существенно снизить объем компрессированных данных при заданном качестве восстановленного кадра.

#### Постановка задачи и перечисление подзадач

Анализ и компенсация движения применяются во всех современных видеокодеках, обеспечивая устранение временной избыточности. В качестве примера на рис. 1 приведена общая схема вейвлет-видеокодера.

Практической целью данной работы является разработка нового метода анализа движения для блока «Анализ движения» и разработка соответствующего метода компенсации движения для блока «Компенсация движения». Для сравнения эффективности предложенного метода с традиционным методом также реализован традиционный метод СБ. Нужно отметить, что для предложенного метода, кроме векторов движения, нужно дополнительно передавать параметры функции аппроксимации для каждого блока.

Для решения поставленной задачи, в работе решены следующие подзадачи:

- Поиск векторов движения традиционным методом СБ с точностью до четверти пиксела и предложенным методом с использованием различных функций аппроксимации. При этом яркость пиксела, смещенного на нецелый вектор движения, определяется с помощью восьмиточечной линейной интерполяции, используемой в стандарте H.265 для интерполяции яркости (H265L) [6, 7]. Параметры метода – точность вектора движения, размер окна поиска, размер блока.

- Компенсация движения методом сопоставления перекрывающихся блоков (СПБ), используемым в вейвлет-видеокодеке Digac [4, 8]. Для адаптации пред-

ложенного метода также модифицирован метод СПБ.

- Получение остатка предсказания, квантованного остатка предсказания (КОП) и восстановленного кадра. Для получения КОП использовано широко распространенное двухканальное вейвлет-преобразование и равномерное квантование с шагом  $Q$ , которые также заданы как параметры.

- Вычисление метрики PSNR (peak signal-to-noise ratio, пиковое отношение сигнал/шум) восстановленного кадра и энтропии КОП для заданного шага квантования  $Q$ . Суммарная энтропия состоит из энтропии КОП и энтропии векторов движения для традиционного метода, а для предложенного метода добавляется энтропия дополнительных параметров, передаваемых с требуемой точностью. Метрика PSNR показывает степень искажения восстановленного кадра по отношению к исходному, а суммарная энтропия оценивает количество битов данных, необходимых для передачи.

- Построение кривой зависимости уровня искажений от скорости выходного потока (rate-distortion curve, далее RD-кривая) на основе PSNR восстановленного кадра и суммарной энтропии по нескольким значениям  $Q$ .

- Построение RD-кривой для нескольких популярных эталонных размеров блоков (8×8 и 16×16).

- Сравнение полученных RD-кривых для традиционного метода СБ и предложенного метода и выводы об эффективности и применимости предложенного метода для задачи сжатия видео.

#### Предложенный метод с использованием различных функций аппроксимации

В данной работе, кроме простой функции аппроксимации  $Y = X$  для улучшения качества предсказания рассмотрены другие элементарные функции, представленные в табл. 1.

Функция 0 используется в традиционном методе, который служит эталонным при сравнении исследованных методов. Качество методов оценивается с помощью RD-кривых. В отличие от традиционного метода, в пред-

ложенном методе для каждого блока, кроме вектора движения, передаются еще дополнительные параметры  $(a, b)$ , которые определяются вместе с вектором движения условием минимизации функции стоимости:

$$SAD_F(V_x, V_y) = \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} |Y(x+m, y+n) - F(X(x+V_x+m, y+V_y+n))|. \quad (2)$$

Здесь  $Y(u, v)$  – значение яркости пикселя с координатами  $(u, v)$  в текущем кадре;  $X(u, v)$  – значение яркости пикселя с координатами  $(u, v)$  в опорном кадре;  $M \times N$  – размер блока;  $(V_x, V_y)$  – вектор движения;  $(x, y)$  – координаты верхнего левого угла рассматриваемого блока в текущем кадре.

Для нахождения параметров  $(a, b)$  можно использовать метод наименьших квадратов [9] или метод ре-

грессии [10]. Нужно отметить, что для реализации предложенного метода нужно решить подзадачи сложности вычислений и способа передачи параметров.

Таблица 1. Список исследованных функций аппроксимации

Номер функции	Функции аппроксимации	Обозначение на графике
0	$F(X) = X$	Ref
1	$F(X) = aX + b$	Linear
2	$F(X) = bX^a$	Power
3	$F(X) = be^{aX}$	Exp
4	$F(X) = a \cdot \ln(X) + b$	Log
5	$F(X) = (a/X) + b$	hyp1
6	$F(X) = X/(bX + a)$	hyp2
7	$F(X) = bX^{aX}$	powOfLinear
8	$F(X) = bX^{a/X}$	powOfHyp

RD-кривые в режиме разбиения кадра на блоки 16x16 для видео "1080p\_riverbed"

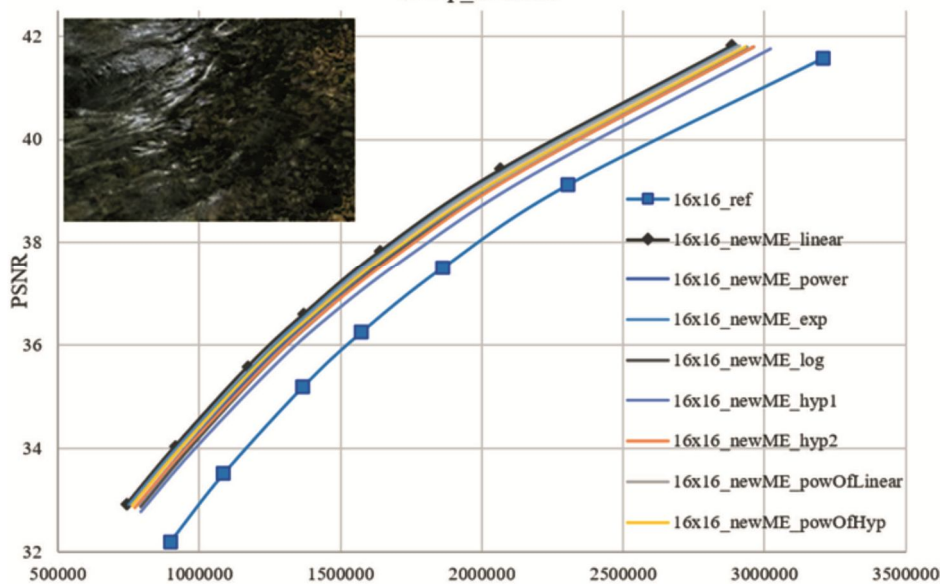


Рис. 2. Сравнение функций аппроксимации для видео «1080p\_riverbed» и размера блока 16×16

RD-кривые в режиме разбиения кадра на блоки 16x16 для видео "1080p\_rush\_hour"

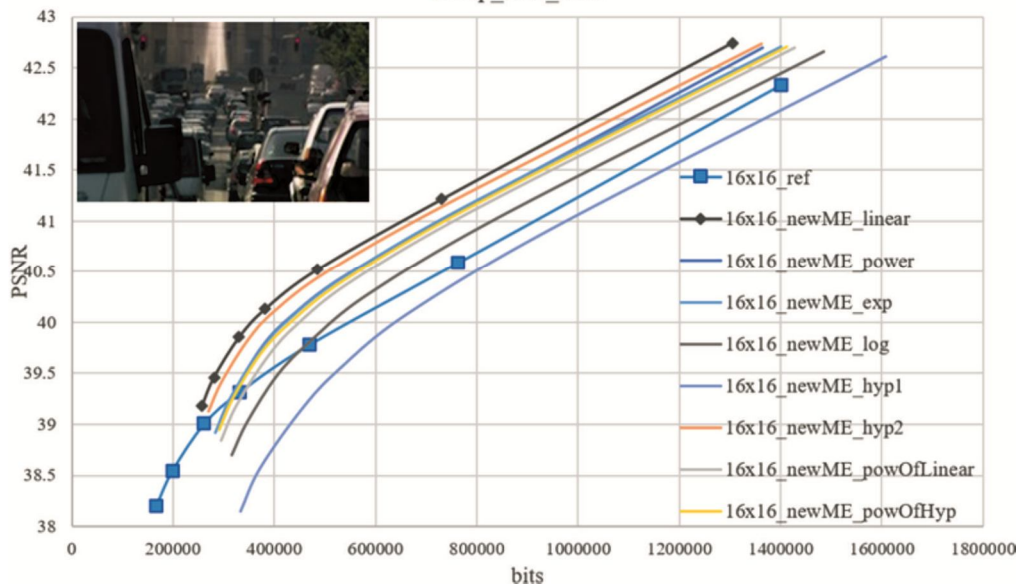


Рис. 3. Сравнение функций аппроксимации для видео «1080p\_rush\_hour» и размера блока 16×16

## Результаты

Проведенный анализ с обработкой достаточно сложных для традиционного метода компенсации движения видеопоследовательностей «1080p\_riverbed» и «1080p\_rush\_hour» [11] показал, что исследованные методы выигрывают у эталонного метода (см. рис. 2 и рис. 3). Можно также сделать вывод, что среди исследованных функций аппроксимации оптимальной является линейная функция (см. номер метода 1 в табл. 1), применение которой для случая размера блока 16×16 позволит снизить объем данных на 15 % для видео «1080p\_riverbed» (см. рис. 2) и 34 % для видео «1080p\_rush\_hour» при достаточно высоком качестве восстановленного кадра (PSNR = 40 дБ).

## Заключение

В работе предложен и разработан новый метод анализа и компенсации движения с использованием различных функций аппроксимации. На основе результатов обработки для дальнейшего использования предложено использовать линейную функцию аппроксимации  $F(X) = aX + b$ . Применение предложенного метода для видео высокого разрешения 1080 p позволяет заметно уменьшить объем данных при заданном качестве восстановленного кадра (от 15 % до 34 % объема данных для размера блока 16×16).

## Литература

1. Дворкович В.П., Дворкович А.В. Цифровые видеотелекоммуникационные системы (теория и практика). Москва: Техносфера, 2012. – 1008 с. ISBN 978-5-94836-336-3.
2. x.264 video codec // <http://www.videolan.org/developers/x264.html>
3. x.265 video codec // <http://x265.org/>
4. Dirac video codec // <https://sourceforge.net/projects/dirac/>
5. ITU-T Recommendation H.265, High efficiency video coding, 04/2013 // ISO/IEC FDIS 23008-2, Information technology – High efficiency coding and media delivery in heterogeneous environments – Part 2: High efficiency video coding (MPEG-H HEVC).
6. Gary J. Sullivan, Jens-Rainer Ohm, Woo-Jin Han, Thomas Wiegand, Overview of the High Efficiency Video Coding (HEVC) Standard // IEEE Transactions on Circuits and Systems for Video Technology, Vol. 22, #12, 12/2012, pp. 1649-1668.
7. Jill M. Boyce, Yan Ye, Jianle Chen, Adarsh K. Ramasubramanian. Overview of SHVC: Scalable Extensions of the High Efficiency Video Coding Standard. IEEE Trans. Circuits Syst. Video Techn. 26(1): 20-34 (2016).
8. Jianguo Zhang Ling Shao Lei Zhang Graeme A. Jones. Intelligent video event analysis and understanding. Springer, 2010 edition. – 251 p. ISBN 978-3-642-17554-1.
9. Гриценко В.А., Белосевич Е.В., Артищева Е.К. Математические методы в географии: Учебное пособие / Калинингр. ун-т. – Калининград, 1999. – 75 с. – ISBN 5-88874-151-5.
10. Фёрстер Э., Рёнц Б. Методы корреляционного и регрессионного анализа. Руководство для экономистов. Перевод с немецкого и предисловие В. М. Ивановой, М.: «Финансы и статистика», 1983 г.- 304 с.
11. Тестовые видео // <https://media.xiph.org/video/derf/>

## ОСОБЕННОСТИ РЕАЛИЗАЦИИ МОДИФИЦИРОВАННОГО МЕТОДА КОМПЕНСАЦИИ ДВИЖЕНИЯ

*Дам Чонг Нам, аспирант Московского физико-технического института (национального исследовательского университета), e-mail: chong.dam@phystech.edu.*

### IMPLEMENTATION FEATURES OF THE MODIFIED MOTION COMPENSATION METHOD

*Dam Trong Nam*

*This work is dedicated to the research and development of the modified motion compensation method for the wavelet-based video codec Dirac. The paper briefly mentions the problem and the essence of the modified method. Then the implementation features of the modified method and the modified method processing results are described. Compared with the traditional motion compensation method, the usage of the modified method in the considered video codec can reduce bitrate up to 20 % for good reconstruction quality PSNR = 40 dB and up to 33 % for excellent quality PSNR = 42 dB for frame splitting case 16x16. In the case of block size 32x32, the proposed method can reduce bitrate from 3 % to 33 % for quality PSNR = 40 dB and from 2 % to 40 % for quality PSNR = 42 dB.*

**Key words:** video coding, approximation function, motion estimation, motion compensation.

**Ключевые слова:** видеокодирование, функция аппроксимации, анализ движения, компенсация движения.

#### Введение

Для уменьшения временной избыточности в видеокodeках используется анализ и компенсация движения. В работе [1] предложен модифицированный метод компенсации движения, основанный на применении другой функции стоимости, в которой введена так называемая функция аппроксимации. Было отмечено, что для применения модифицированного метода нужно решить проблемы, связанные с дополнительными параметрами функции аппроксимации, такие как метод их определения, требуемая точность и способ их передачи. С целью проверки практической значимости модифицированного метода разработана экспериментальная программа, написанная на C++, которая используется для исследования предложенных функций аппроксимации. По оценке энтропии данных, необходимых для кодирования, и по качеству восстановленного кадра был сделан вывод, что среди исследованных функций аппроксимации наиболее подходящей для задачи сжатия видео является линейная функция аппроксимации и модифицированный метод может уменьшить объем кодируемых данных при заданном качестве восстановления.

Данная работа посвящена внедрению модифицированного метода в конкретный видеокodeк Dirac и описанию особенностей модифицированного метода. Настоящая статья организована следующим образом. Сначала коротко приведена постановка задачи, затем описаны особенности модифицированного метода компенсации движения, и в конце статьи приведены результаты обработки по модифицированному методу.

*Работа посвящена исследованию и разработке модифицированного метода компенсации движения для вейвлет-видеокodeка Dirac. В работе коротко упомянуты постановка задачи и суть модифицированного метода. Затем описаны особенности реализации модифицированного метода и результаты обработки по модифицированному методу. По сравнению с традиционным методом компенсации движения, применение модифицированного метода в рассматриваемом видеокodeке может уменьшить выходной поток до 20 % при хорошем качестве восстановления PSNR = 40 дБ и до 33 % при отличном качестве PSNR = 42 дБ для случая разбиения кадра на блоки 16x16. А для случая размера блока 32x32, применение предложенного метода позволит уменьшить выходной поток от 3 % до 33 % для качества PSNR = 40 дБ и от 2 % до 40 % для качества PSNR = 42 дБ.*

#### Постановка задачи

Для уменьшения временной избыточности в современных видеокodeках используются процедуры анализа и компенсации движения. С целью исследования и проверки практической значимости модифицированного метода компенсации движения был выбран вейвлет-видеокodeк Dirac. При реализации предложенного метода кроме векторов движения нужно дополнительно передавать параметры функции аппроксимации для каждого блока.

#### Особенности модифицированного метода компенсации движения

В работе [1] кроме простейшей функции аппроксимации  $Y = X$ , соответствующей традиционному методу компенсации движения, для улучшения качества предсказания рассмотрены другие элементарные функции, представленные в табл. 1.

Традиционный метод служит эталоном для сравнения предложенных методов с более сложной функцией аппроксимации. Качество методов оценивается кривой



зависимости уровня искажений от скорости выходного потока (rate-distortion curve, далее RD-кривая). В модифицированном методе для каждого блока кроме вектора движения передаются еще дополнительные параметры  $(a, b)$  (см. [1]), которые определяются вместе с вектором движения условием минимизации функции стоимости:

$$SAD(V_x, V_y) = \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} |Y(x+m, y+n) - F(X(x+V_x+m, y+V_y+n))| \quad (1)$$

Таблица 1. Список исследованных функций аппроксимации

Номер метода	Функции аппроксимации	Обозначение на графике
0	$F(X) = X$	ref
1	$F(X) = aX + b$	linear
2	$F(X) = bX^a$	power
3	$F(X) = be^{aX}$	exp
4	$F(X) = a \cdot \ln(X) + b$	log
5	$F(X) = (a/X) + b$	hyp1
6	$F(X) = X/(bX + a)$	hyp2
7	$F(X) = bX^{aX}$	powOfLinear
8	$F(X) = bX^{a/X}$	powOfHyp

Здесь  $Y(u, v)$  – значение яркости пикселя с координатами  $(u, v)$  в текущем кадре;  $X(u, v)$  – значение яркости пикселя с координатами  $(u, v)$  в опорном кадре;  $M \times N$  – размер блока;  $(V_x, V_y)$  – вектор движения;  $(x, y)$  – координаты верхнего левого угла рассматриваемого блока в текущем кадре.

**Определение дополнительных параметров**

Дополнительные параметры определяются на этапе поиска движения вместе с вектором движения. К примеру, для первой функции – линейной функции аппроксимации, параметры  $a, b$  можно вычислить, используя метод наименьших квадратов:

$$\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} (Y(x+i, y+j) - (aX(x+i, y+j) + b))^2 \rightarrow Min. \quad (2)$$

Решение условия (2) представлено формулой (3):

$$a = \frac{\overline{XY} - \overline{X}\overline{Y}}{\overline{X^2} - (\overline{X})^2}; \quad b = \overline{Y} - a\overline{X}, \quad (3)$$

где  $\overline{X}$  обозначает среднее значение  $X$  по всем пикселям блока.

Другие функции аппроксимаций могут привести к линейной зависимости путем преобразования переменных, процесс поиска движения выполняется аналогичным образом как для первой функции аппроксимации. К

примеру, для второго режима, сделав преобразование по формуле (4) получим линейную функцию новых переменных:

$$Y = bX^a \xrightarrow{y=\ln(Y), x=\ln(X); A=a, B=\ln(b)} y = Ax + B \quad (4)$$

В этом случае под дополнительными параметрами понимаются параметры  $(A, B)$  полученной линейной зависимости которые аналогично вычисляются по формуле (3) с использованием новых переменных:

$$A = \frac{\overline{xy} - \overline{x}\overline{y}}{\overline{x^2} - (\overline{x})^2}, \quad B = \overline{y} - A\overline{x}. \quad (5)$$

Заметим, что для общности можно записать, что для первой функции аппроксимации выполняется преобразование  $x = X, y = Y, A = a, B = b$ . Таким образом для удобства обработки целесообразно ввести преобразования и новые формулы передаваемых дополнительных параметров, представленные в табл. 2.

С помощью формул преобразования, приведенных в табл. 2, для всех исследованных функций аппроксимации получаются:

– единственная функция аппроксимации в новых переменных:  
 $y = Ax + b;$  (6)

– дополнительные параметры в новых переменных, которые сохраняются/передаются и используются для предсказания кадра с помощью предложенного способа:

$$A = \frac{\overline{xy} - \overline{x}\overline{y}}{\overline{x^2} - (\overline{x})^2}, \quad B = \overline{y} - A\overline{x}. \quad (7)$$

**Точность и способ передачи дополнительных параметров**

По формуле (7) легко заметить, что в общем случае дополнительные параметры для всех исследованных функций аппроксимации принимают действительные, а не целые значения, так что нужно сохранять/передавать дополнительные параметры с необходимой точностью, иначе качество восстановленного кадра будет невысоким.

Нужно отметить, что модифицированный метод компенсации движения ориентируется на высокое качество предсказания, так что сначала необходимо определить высокие точности параметров для каждой функции аппроксимации, которые обеспечивают максимальное качество предсказания и превышение которых практически не дает улучшения в качестве предсказания, а далее исследовать точности дополнительных параметров пу-

Таблица 2. Формулы преобразования переменных и передаваемых параметров исследованных функций аппроксимации

Номер функции аппроксимации	Функции аппроксимации	Формулы преобразования	Формулы передаваемых параметров
1	$Y = aX + b$	$x = X, y = Y$	$A = a, B = b$
2	$Y = bX^a$	$x = \ln(X), y = \ln(Y)$	$A = a, B = \ln(b)$
3	$Y = be^{aX}$	$x = X, y = \ln(Y)$	$A = a, B = \ln(b)$
4	$Y = a \cdot \ln(X) + b$	$x = \ln(X), y = Y$	$A = a, B = b$
5	$Y = (a/X) + b$	$x = 1/X, y = Y$	$A = a, B = b$
6	$Y = X/(bX + a)$	$x = 1/X, y = 1/Y$	$A = a, B = b$
7	$Y = bX^{aX}$	$x = X \cdot \ln(X), y = \ln(Y)$	$A = a, B = \ln(b)$
8	$Y = bX^{a/X}$	$x = \ln(X)/X, y = \ln(Y)$	$A = a, B = \ln(b)$

тем подбора менее высоких точностей и оценивать итоговые результаты при снижении точности дополнительных параметров.

Самая высокая точность дополнительных параметров (A, B) для первой функции можно определить по формуле (8):

$$Y = Y(a, b) = aX + b = AX + B \rightarrow \Delta Y = X\Delta A + \Delta B$$

$$X \in [1, 255] \rightarrow \Delta A \approx 0,01; \Delta B \approx 1 \rightarrow \Delta Y \in [1, 4]$$

С указанной точностью параметров предсказание яркости пикселя осуществляется с ошибкой от 1 до 4, что гарантирует высокое качество предсказания.

Аналогично для второй функции аппроксимации, высокая точность дополнительных параметров (A, B) определяется по формуле 9:

$$Y = bX^a \xrightarrow{y=\ln(Y), x=\ln(X); A=a, B=\ln(b)} y = Ax + B$$

$$y = Ax + B \rightarrow \Delta y = x\Delta A + \Delta B$$

$$\xrightarrow{\Delta y = \frac{\Delta Y}{Y}; x = \ln(X)} \Delta Y = Y \ln(X)\Delta A + Y\Delta B$$

$$X, Y \in [1, 255] \rightarrow \Delta A \approx 0,001; \Delta B \approx 0,01$$

Выполнив аналогичные оценки для всех предложенных функций аппроксимации, получим самые высокие точности дополнительных параметров для исследованных функций аппроксимации, представленные в табл. 3.

#### Способ передачи дополнительных параметров

Как и вектор движения, каждый компонент A и B передается отдельно в связи с тем, что каждый компонент

принимает значения из определенного диапазона и эти значения оказываются в какой-то степени коррелированы.

#### Результаты

В работе кадр разбивается на блоки некоторых размеров, таких как, 32×32, 16×16, 8×8. В качестве примера на рис. 1-2 представлены RD-кривые исследованных функций аппроксимации для видео «1080p\_riverbed» и «1080p\_rush\_hour» [5] в случае размера блока 16×16.

Из рис. 1-2 можно сделать вывод, что среди исследованных функций аппроксимации наиболее подходящими являются функции 1, 2, 6 (см. табл. 1). При этом наилучшей является первая функция – линейная функция аппроксимации. Если сравнить с традиционным методом поиска и компенсации движения, результат которого обозначается «16x16\_ref» на рис. 1-2, то модифицированный метод с различными функциями аппроксимации выигрывает у традиционного метода. Сделанный выше вывод также подтверждается аналогичным анализом для разных размеров блока с рядом видео различных типов [5].

Если в качестве функции аппроксимации выбрана наилучшая функция – линейная и проводится сравнение модифицированного метода с традиционным для разных размеров блока, то получаются результаты, представленные на рис. 3-6 в качестве примера для видео «1080p\_riverbed» и «1080p\_rush\_hour».

Таблица 3. Оценка самых высоких точностей параметров исследованных функций аппроксимации для задачи сжатия видео

№	Формула функций аппроксимации	Оценка ошибки квантования передаваемых параметров	Самые высокие нужные точности параметров
1	$F(X) = aX + b$	$\Delta Y = X\Delta A + \Delta B$	$\Delta A \approx 0,01; \Delta B \approx 1$
2	$F(X) = bX^a$	$\Delta Y = Y \ln(X)\Delta A + Y\Delta B$	$\Delta A \approx 0,001; \Delta B \approx 0,01$
3	$F(X) = beaX$	$\Delta Y = YX\Delta A + Y\Delta B$	$\Delta A \approx 0,0001; \Delta B \approx 0,01$
4	$F(X) = a \cdot \ln(X) + b$	$\Delta Y = \ln(X)\Delta A + \Delta B$	$\Delta A \approx 0,1; \Delta B \approx 1$
5	$F(X) = (a/X) + b$	$\Delta Y = (1/X)\Delta A + \Delta B$	$\Delta A \approx 10; \Delta B \approx 1$
6	$F(X) = X/(bX + a)$	$\Delta Y = (Y^2 / X)\Delta A + Y^2\Delta B$	$\Delta A \approx 0,01; \Delta B \approx 0,0001$
7	$F(X) = bX^{aX}$	$\Delta Y = YX \ln(X)\Delta A + Y\Delta B$	$\Delta A \approx 0,00001; \Delta B \approx 0,01$
8	$F(X) = bX^{a/X}$	$\Delta Y = (Y \ln(X) / X)\Delta A + Y\Delta B$	$\Delta A \approx 0,1; \Delta B \approx 0,01$

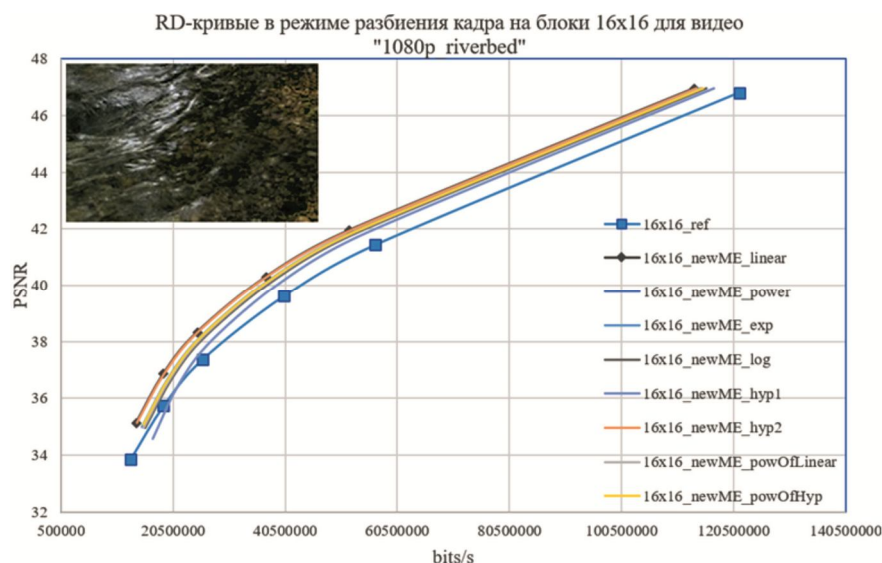


Рис. 1. RD-кривые в режиме разбиения кадра на блоки 16×16 для видео «1080p\_riverbed»

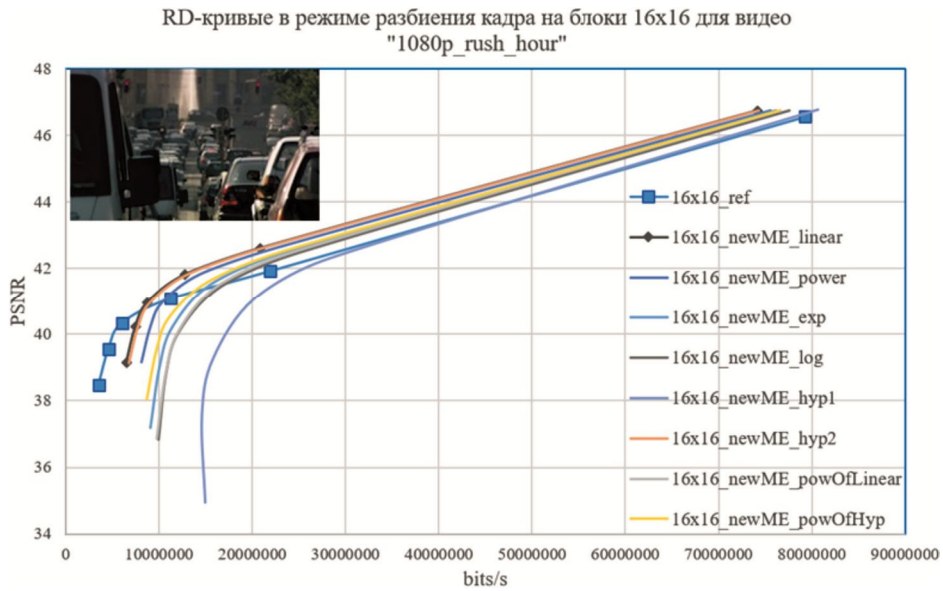


Рис. 2. RD-кривые в режиме разбиения кадра на блоки 16×16 для видео «1080p\_rush\_hour»

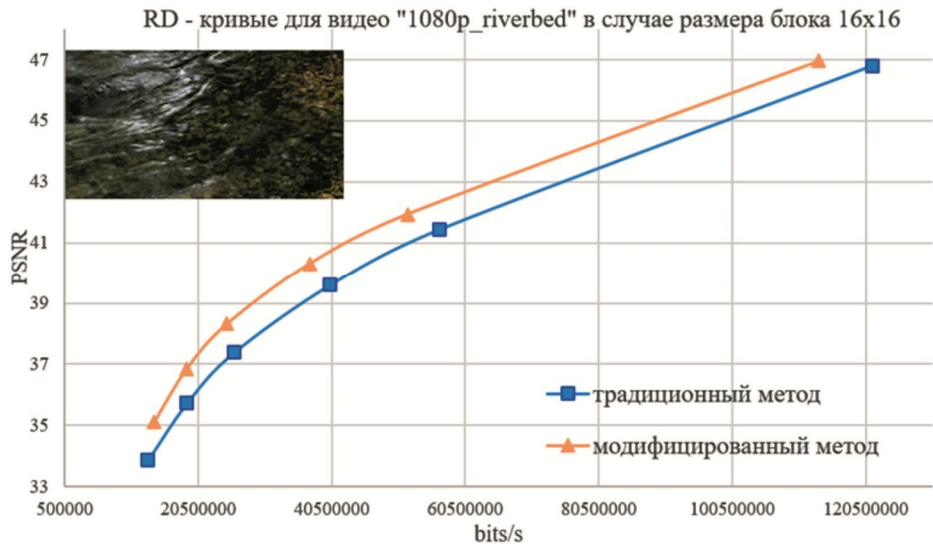


Рис. 3. Сравнение RD-кривых модифицированного и традиционного методов для видео «1080p\_riverbed» в случае размера блока 16×16

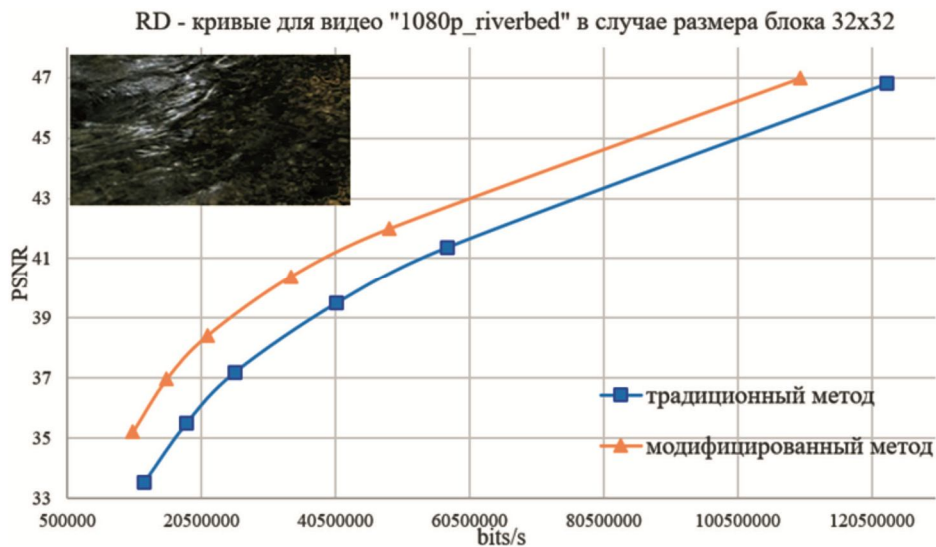


Рис. 4. Сравнение RD-кривых модифицированного и традиционного методов для видео «1080p\_riverbed» в случае размера блока 32×32

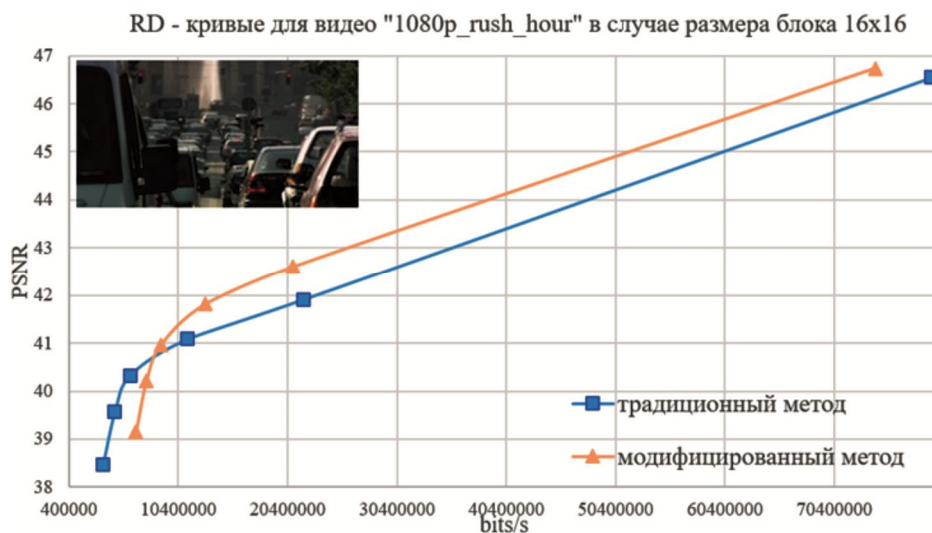


Рис. 5. Сравнение RD-кривых модифицированного и традиционного методов для видео «1080p\_rush\_hour» в случае размера блока 16×16

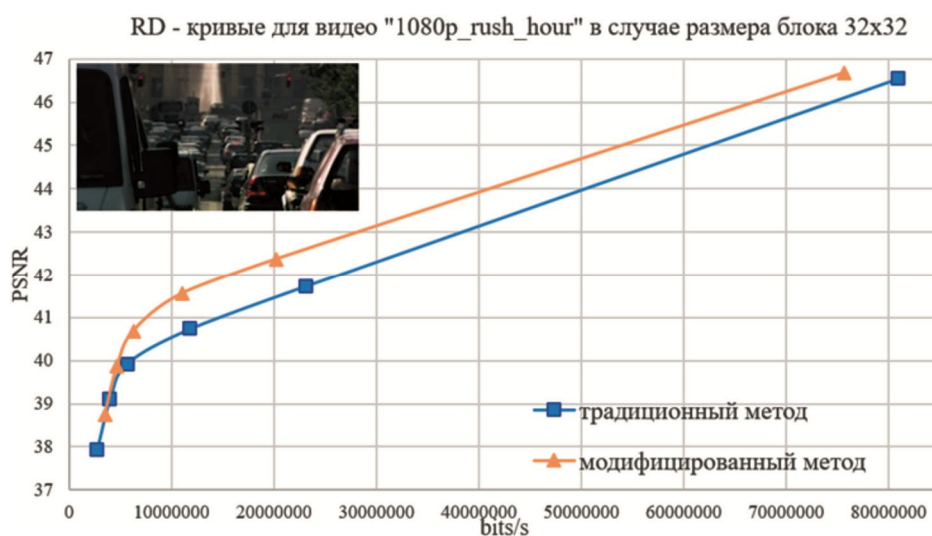


Рис. 6. Сравнение RD-кривых модифицированного и традиционного методов для видео «1080p\_rush\_hour» в случае размера блока 32×32

Таблица 4. Выигрыш модифицированного метода в сравнении с традиционным методом по скорости выходного потока при заданном качестве восстановления

видео \ PSNR	16×16		32×32	
	40	42	40	42
1080p_blue_sky	-	2 %	5,5 %	9 %
1080p_pedestrian_area	10 %	20 %	33 %	29 %
1080p_riverbed	19,5 %	17 %	30 %	25 %
1080p_rush_hour	-	33 %	20 %	40 %
1080p_station2	0	12,5 %	19 %	15 %
1080p_sunflower	-	9 %	12,5 %	12 %
1080p_tractor	3 %	7,5 %	17 %	14 %
720p_mobcal_ter	2 %	2 %	4,5%	2,5 %
720p_shields_ter	2 %	4 %	3 %	2 %
720p_stockholm	3 %	2 %	4 %	3 %
4cif_crew	17 %	14 %	22 %	17 %
4cif_soccer	4 %	4 %	8,5%	7 %

Проведение численного сравнения скорости выходного потока при заданном качестве восстановления для

ряда видео различных типов [5] в разных случаях размера блока, привело к результатам, представленные в табл. 4.

Из табл. 4 и также из графиков, построенных при проведенном анализе, можно сделать следующие выводы:

– Модифицированный метод выигрывает у традиционного метода при хорошем качестве восстановления (например, PSNR = 40) и при высоком качестве восстановления (например, PSNR = 42). В общем случае RD-кривые для предложенного метода и эталонного метода пересекаются в точке со значением PSNR в районе от 35 до 37 дБ, то есть, для участка RD-кривой с относительно низким качеством модифицированный метод проигрывает традиционному, однако этот участок не является интересующим для многих приложений сжатия видео.

– Для видео высокого разрешения 1080 p (1920x1080) модифицированный метод существенно уменьшает выходной поток при заданном качестве восстановления, что также объясняет актуальность и практическую значимость модифицированного метода в связи с тем, что требование пользователей увеличивается и часто нужно передавать видео высокого разрешения с требованием высокого качества восстановления.

– Чем сложнее тестовые видео для предсказания (изменения в сцене существенно отличаются от плоскопараллельного перемещения), тем лучше модифицированный метод по сравнению с традиционным.

– Для случая размера блока 16×16 модифицированный метод может сократить выходной поток до 20 % при хорошем качестве восстановления PSNR = 40 дБ и до 33 % при отличном качестве восстановления PSNR = 42 дБ. А для случая размера блок 32×32 применение модифицированного метода позволяет уменьшить выходной поток от 3% до 33% для качества восстановления PSNR = 40 дБ и от 2% до 40% для качества восстановления PSNR = 42 дБ.

## Заключение

В работе коротко приведена постановка задачи, состоящая в том, что нужно адаптировать в вейвлет-видеокодек Dirac модифицированный метод компенсации движения, основанный на новой функции стоимости, в которой введена функция аппроксимации. В работе также описаны особенности модифицированного метода и результаты обработки по модифицированному методу. Модифицированный метод внедрен в вейвлет-видеокодек Dirac и по результатам обработки множественных видео различных типов получены следующие выводы:

1. Применение модифицированного метода в рассматриваемом видеокодеке может уменьшить выходной поток до 33 % при хорошем качестве восстановления PSNR = 40 дБ и до 40 % при отличном качестве PSNR = 42 дБ.

2. Чем сложнее тестовые видео для предсказания (изменения в сцене отличаются от плоскопараллельного перемещения), тем лучше модифицированный метод по сравнению с традиционным.

## Литература

1. Дам Чонг Нам. Улучшение качества традиционного метода компенсации движения // Цифровая обработка сигналов. 2019. №3 (настоящий номер).
2. Dirac video codec // <https://sourceforge.net/projects/dirac/>
3. Гриценко В.А., Белосевич Е.В., Артищева Е.К. Математические методы в географии: Учебное пособие / Калинингр. ун-т. – Калининград, 1999. – 75 с. – ISBN 5-88874-151-5.
4. Фёрстер Э., Рёнц Б. Методы корреляционного и регрессионного анализа. Руководство для экономистов. Перевод с немецкого и предисловие В. М. Ивановой, М.: «Финансы и статистика», 1983 г. – 304 с.
5. Тестовые видео // <https://media.xiph.org/video/derf/>

## МЕТОДЫ И АЛГОРИТМЫ ОБНАРУЖЕНИЯ ПАУЗ В РЕЧИ

*Волченков В.А., старший преподаватель кафедры телекоммуникаций и основ радиотехники  
ФГБОУ ВО «Рязанского государственного радиотехнического университета им. В.Ф. Уткина»,  
e-mail: volchenkov.rzn@yandex.ru.*

## VOICE ACTIVITY DETECTION METHODS AND ALGORITHMS

*Volchenkov V.A.*

*Problems of accuracy increase in voice activity detection (VAD) are considered. General information about some VAD methods is given. A new method of voice activity detection is offered. Performance comparison of VADs is given.*

**Key words:** voice activity detection, problems of accuracy, performance comparison.

**Ключевые слова:** обнаружение пауз, детектор активности речи, методы и алгоритмы, сравнительный анализ.

### Введение

Реализация существующих алгоритмов обнаружения пауз базируется на предположении, что речь является нестационарным сигналом, форма ее спектра обычно изменяется через короткие отрезки времени – 10-30 мс. Также считают, что фоновый шум обычно стационарен на более длинном отрезке времени, немного изменяясь со временем, а уровень речевого сигнала обычно выше уровня фонового шума [1]. Речь обычно делят на отрезки длительностью 16-32 мс, и анализируют уровень энергии сигнала на каждом интервале, а также количество переходов сигнала через ноль. В том случае, когда временной интервал определяется обнаружителем как пауза, перед окончательным принятием решения, что сигнал отсутствует, системе необходимо последовательно продетектировать ещё несколько фреймов (в системе GSM 5-6). Таким образом, существующие на сегодняшний день способы определения активности речи позволяют выявить паузы, длительность которых значительно превышает 40 мс. Обнаружение коротких пауз и установление более точных границ для длинных пауз становятся важными задачами, решению которых и посвящена данная работа.

В настоящей работе предложен детектор активности речи, обеспечивающий существенное повышение вероятности правильного разделения речевых сигналов на периоды активной речи и паузы.

### Сравнение эффективности алгоритмов VAD

Проведем сравнение эффективности предлагаемого способа с алгоритмом VAD кодера G.729B, а также со способом детектирования пауз на основе отношения правдоподобия (VAD LR – Likelihood-Ratio-Based VAD).

#### Алгоритм VAD кодера G.729B

Кодер G.729B делит речь на интервалы по 10 мс и вырабатывает решение о наличии или отсутствии речи для каждого фрейма, оценивая при этом четыре параметра [2, 3, 4]:

*Рассмотрены вопросы увеличения точности обнаружения пауз. Приведена общая информация о некоторых методах детектирования активности речи. Представлен новый способ обнаружения пауз в речи. Приведено сравнение их производительности.*

- разность энергий всего диапазона –  $\Delta E_f = \overline{E_f} - E_f$ ,
- разность энергий диапазона НЧ –  $\Delta E_l = \overline{E_l} - E_l$ ,
- искажение спектра –  $\Delta LSF = \sum_{i=0}^9 (\overline{LSF_i} - LSF_i)^2$ ,
- разность частоты переходов через ноль –  $\Delta ZC = \overline{ZC} - ZC$ , где  $E_f$  – энергия всего диапазона,  $E_l$  – энергия диапазона НЧ,  $LSF_i$  –  $i$ -я частота спектра сигнала и  $ZC$  – частота переходов через ноль входного сигнала,  $\overline{E_f}$ ,  $\overline{E_l}$ ,  $\overline{LSF_i}$ ,  $\overline{ZC}$  – параметры, характеризующие шум и обновляемые посредством анализа фонового шума.

Блок схема алгоритма VAD кодера G.729B представлена на рис. 1 [4]. Входные параметры для анализа VAD могут быть получены из входного сигнала или из промежуточных значений речевого кодера. Затем рассчитываются параметры разницы между параметрами входного сигнала и шума  $\Delta E_f$ ,  $\Delta E_l$ ,  $\Delta LSF$  и  $\Delta ZC$ . Решение о наличии речи получают путем анализа интервалов речевого сигнала по четырем параметрам, которые поступают на схему анализа предыдущих решений. Блок обновления параметров шума основан на схеме авторегрессии первого порядка. Они обновляются, если разница энергии всего диапазона меньше заданного фиксированного порога.

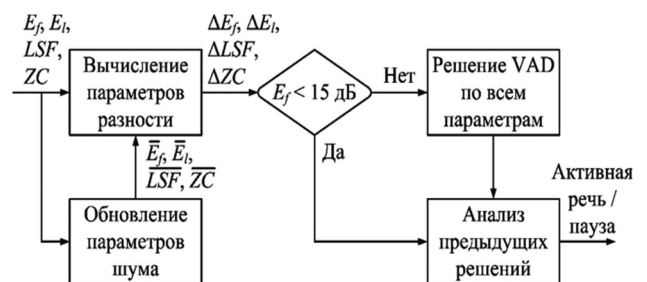


Рис. 1. Алгоритм VAD кодера G.729B

**Алгоритм VAD на основе отношения правдоподобия**

В [5] был предложен метод, который в отличие от традиционных методов VAD базируется на статистической модели [2], и, как описано, позволяет значительно повысить точность обнаружения [5]. Причиной высокой производительности является учет критериев подавления шума, предложенных Ефремом и Малом [6] для критериев принятия решения о речевой активности.

Решение о речевой активности может быть рассмотрено, как проверка гипотез:  $H_0$  и  $H_1$ , которые означают отсутствие и присутствие речи, соответственно. Предполагая, что каждая спектральная составляющая речи и шума имеет комплексное гауссово распределение [6], в котором шум является аддитивным и не коррелирован с речью, функции плотности условного распределения вероятностей (conditional probability density functions – PDF) спектральной составляющей шума  $Y_k$ , данных  $H_{0,k}$  и  $H_{1,k}$  запишем следующим образом:

$$p(Y_k | H_{0,k}) = \frac{1}{\pi \lambda_{N,k}} \exp \left\{ -\frac{|Y_k|^2}{\lambda_{N,k}} \right\}, \quad (1)$$

$$p(Y_k | H_{1,k}) = \frac{1}{\pi (\lambda_{N,k} + \lambda_{X,k})} \exp \left\{ -\frac{|Y_k|^2}{\lambda_{N,k} + \lambda_{X,k}} \right\}, \quad (2)$$

где  $k$  означает индекс элемента спектральной выборки, а  $\lambda_{N,k}$  и  $\lambda_{X,k}$  обозначают дисперсию спектров шума и речи соответственно.

Отношение правдоподобия (likelihood ratio – LR)  $k$  – го элемента спектральной выборки  $\Lambda_k$  определяется из упомянутых выше двух функций плотности условного распределения вероятностей [5]:

$$\Lambda_k = \frac{p(Y_k | H_{1,k})}{p(Y_k | H_{0,k})} = \frac{1}{1 + \xi_k} \exp \left\{ \frac{(1 + \gamma_k) \xi_k}{1 + \xi_k} \right\}, \quad (3)$$

где  $\gamma_k$  и  $\xi_k$  являются апостериорным и априорным ОСШ, определяемые как  $\gamma_k = |Y_k|^2 / \lambda_{N,k} - 1$  и  $\xi_k = \lambda_{X,k} / \lambda_{N,k}$ . Заметим, что определение апостериорного ОСШ немного отличается от оригинального,  $\gamma_k = |Y_k|^2 / \lambda_{N,k}$  [7]. Предполагается, что дисперсия шума известна в результате адаптации шума. Тем не менее, дисперсия речи является неизвестной, таким образом, априорное ОСШ  $n$ -го фрейма  $\xi_k^{(n)}$  оценивают, используя метод прямого принятия решения (decision-directed (DD) method) [6]:

$$\hat{\xi}_k^{(n)} = \alpha \frac{|\hat{X}_k^{(n-1)}|^2}{\lambda_{N,k}^{(n-1)}} + (1 - \alpha) \text{MAX} \{ \gamma_k^{(n)}, 0 \}, \quad (4)$$

где  $\alpha$  – взвешивающий элемент, например, 0,98, и амплитуда спектра незашумленной речи  $|\hat{X}_k|$  оценивается с использованием минимальной среднеквадратической ошибки оценочной функции амплитуды логарифмического спектра [7]. Решение о наличии речевой активности представляется средним геометрическим значений  $\Lambda_k$  для всех спектральных выборок:

$$\Lambda = \exp \left\{ \frac{1}{K} \sum_{k=1}^K \log \Lambda_k \right\}, \quad (5)$$

где  $K$  обозначает количество спектральных выборок.

Апостериорное ОСШ  $\gamma_k$  сильно колеблется от фрейма к фрейму из-за большой флуктуации амплитуды

спектра на кратковременном интервале  $|Y_k|$ . С другой стороны, априорное ОСШ  $\hat{\xi}_k$  меняется медленно вследствие сглаживающего эффекта. Если значение  $\alpha$  увеличивается,  $\hat{\xi}_k$  становится более сглаженной. Изменения  $\gamma_k$  и  $\hat{\xi}_k$  уравнивают друг друга при вычислении  $\Lambda_k$  и, следовательно, в результате увеличивается производительность VAD. Потому оценка DD для априорного ОСШ полезна не только для избегания явления музыкального шума при усилении речи [8], но также для уменьшения количества ошибок в обнаружении речевой активности.

**Предлагаемый способ детектирования активности речи**

В настоящей работе предложен детектор активности речи, базирующийся на использовании вспомогательного сигнала специальной частоты [10].

Структурная схема детектора активности речи изображена на рис. 2.

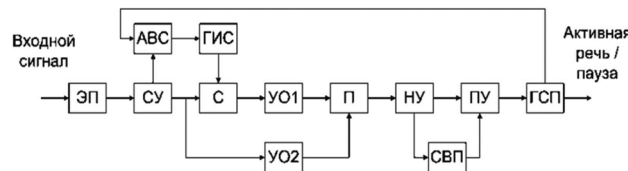


Рис. 2. Структурная схема детектора активности речи

В состав устройства входят: электроакустический преобразователь (ЭП), селективный усилитель (СУ), анализатор входного сигнала (АВС), генератор измерительного сигнала (ГИС), сумматор (С), амплитудные усилители-ограничители (УО) 1 и 2, перемножитель (П), накопитель-усреднитель (НУ), схема вычисления порога (СВП) и пороговое устройство (ПУ) и генератор сигнала паузы (ГСП).

Речевой сигнал с выхода электроакустического преобразователя усиливается селективным усилителем и подается на анализатор входного сигнала и вход сумматора. На второй вход сумматора подается сигнал с выхода генератора измерительного сигнала. Значение необходимой мощности измерительного сигнала вычисляется в анализаторе входного сигнала путем анализа первых 100 мс, т.к. на этом интервале речь обычно отсутствует, и пересчитывается заново при принятии системой решений о наличии паузы в течение 500 мс подряд. Суммарный сигнал с выхода сумматора поступает на вход усилителя-ограничителя 1, где происходит усиление сигнала, а затем ограничение по амплитуде. Аналогичная операция проводится над сигналом, поступающим с выхода селективного усилителя на вход усилителя-ограничителя 2. Сигнал с выхода усилителя-ограничителя 1 подается на первый вход перемножителя. На второй вход перемножителя подается сигнал с выхода усилителя-ограничителя 2. Сигнал с выхода перемножителя поступает на вход накопителя-усреднителя, где происходит выделение сигнала, по амплитуде которого принимают решение о наличии периода активного речевого сигнала или паузы в пороговом устройстве. Значение порога вычисляется в схеме вычисления порога пу-



тем анализа интервала длительностью 50 мс после изменения мощности вспомогательного сигнала. Сигнал с выхода порогового устройства поступает на вход генератора сигнала паузы, который при наличии паузы генерирует сигнал паузы длительностью 10 мс.

### Сравнение эффективности предлагаемого способа с методами обнаружения пауз

Тестовым сигналом была речь длительностью 108 секунд, состоящая из фраз, надиктованных разными дикторами, смешанная с транспортным шумом с ОСШ: 5, 10, 15, 20 и 25 дБ. Интервалы активной речи и паузы были отмечены вручную. Пропорции между неактивными и активными участками речи были 0,46 и 0,54, соответственно. Результаты оценки ошибок определения речи и пауз для приведенных выше методов представлены в табл. 1-3.

Таблица 1. Оценка ошибок определения речи и пауз алгоритмом VAD кодера G.729B

ОСШ, дБ	Ошибка определения речи, %	Ошибка определения пауз, %
5	19,16	1,71
10	12,58	3,97
15	5,34	5,89
20	4,57	7,71
25	2,62	10,27

Таблица 2. Оценка ошибок определения речи и пауз способом детектирования пауз на основе отношения правдоподобия

ОСШ, дБ	Ошибка определения речи, %	Ошибка определения пауз, %
5	9,73	20,73
10	3,13	18,52
15	1,67	15,47
20	0,89	10,29
25	0,42	9,58

Таблица 3. Оценка ошибок определения речи и пауз разрабатываемым способом детектирования пауз

ОСШ, дБ	Ошибка определения речи, %	Ошибка определения пауз, %
5	21,47	2,41
10	9,71	2,53
15	5,01	3,09
20	2,03	4,14
25	0,29	6,86

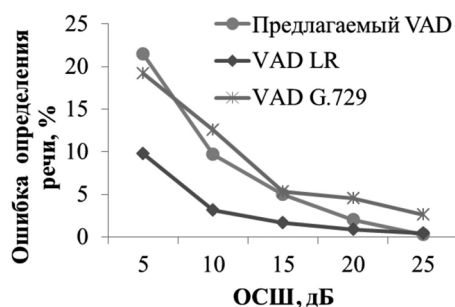


Рис. 3. Сравнение производительности методов обнаружения пауз в речи при воздействии на речевой сигнал транспортного шума. Ошибка определения речи по отношению к ОСШ

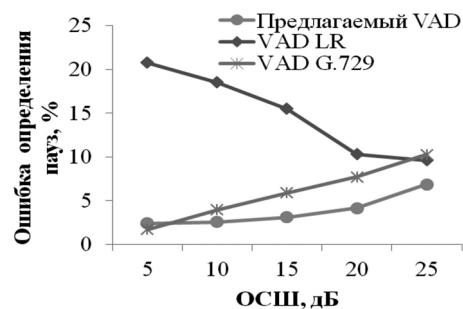


Рис. 4. Сравнение производительности методов обнаружения пауз в речи при воздействии на речевой сигнал транспортного шума. Ошибка определения пауз по отношению к ОСШ

### Заключение

Предложенный способ показал себя лучше алгоритма VAD кодера G.729B почти для всех значений ОСШ. Только при ОСШ равном 5 дБ алгоритм VAD кодера G.729B справился лучше. Способ детектирования пауз на основе отношения правдоподобия показал себя лучше остальных приведенных способов с точки зрения параметра «ошибка определения речи», но с точки зрения параметра «ошибка определения пауз» оказался значительно хуже остальных. Самую низкую ошибку определения пауз почти для всех значений ОСШ показал способ детектирования пауз, предложенный в данной статье. В дальнейшей работе предполагается улучшить разрабатываемый алгоритм для уменьшения ошибки определения речи.

### Литература

1. Шелухин О.И. Цифровая обработка и передача речи / О.И. Шелухин, В.Г. Лукьянцев; Под ред. О.И. Шелухина. – М.: Радио и связь, 2000. – 456 с.: ил.
2. Kondoz A.M. Digital Speech. Coding for Low Bit Rate Communication Systems. – John Wiley & Sons, Ltd. 2004. – 442 p.
3. ITU-T (1996) A silence compression scheme for G.729 optimised for terminals conforming to ITU-T V.70, ITU-T Rec. G.729 Annex B.
4. ITU-T (1996) Coding of speech at 8 kbit/s using conjugate-structure algebraiccode excited linear prediction (CS-ACELP), ITU-T Rec. G.729.
5. Sohn J., Kim N.S., and Sung W. (1999) 'A statistical model-based voice activity detection', in IEEE Signal Processing Letters, 6(1):1-3.
6. Y. Ephraim and D. Malah (1984) 'Speech enhancement using a minimum mean square error short-time spectral amplitude estimator', in IEEE Trans. on Acoust., Speech and Signal Processing, 32(6):1109-20.
7. Y. Ephraim and D. Malah (1985) 'Speech enhancement using a minimum mean square error log-spectral amplitude estimator', in IEEE Trans. on Acoust., Speech and Signal Processing, 33(2):443-5.
8. O. Capp'e (1994) 'Elimination of musical noise phenomenon with the Ephraim and Malah noise suppression', in IEEE Trans. Speech and Audio Processing, 2(2):345-9.
9. Пат. 2436173 Российская Федерация, МПК G10L 15/00, G10L 11/02, Способ обнаружения пауз в речевых сигналах и устройство его реализующее / Витязев В.В., Розов В.И., Волченков В.А.; заявитель и патентообладатель Рязанский государственный радиотехнический университет. – № 2010124342/08, заяв. 15.06.10; опубл. 10.12.11, Бюл. 34.



УДК 621.397

## ОПТИЧЕСКИЕ ЧАСТОТНО-КОНТРАСТНЫЕ ХАРАКТЕРИСТИКИ АТМОСФЕРЫ

*Сорока Е.З., к.т.н., гл. научный сотрудник ЗАО «МНИТИ» (Московский научно-исследовательский телевизионный институт), e-mail: soroka@mniti.ru;*

*Филатов В.С., к.т.н., ведущий научный сотрудник ЗАО «МНИТИ» (Московский научно-исследовательский телевизионный институт), e-mail: filatov\_vs@mniti.ru.*

## THE OPTICAL MODULATION-TRANSFER FUNCTION OF THE ATMOSPHERE

*Soroka E.Z., Filatov V.S.*

*The quality of images obtained in optical vision and registration systems (in vision, UF and IR ranges) depends not only on these systems parameters' but on the atmospheric image transmission channel parameters. The effects of an atmospheric channel consist of overall optical signal attenuation and background illumination but also in high frequency components attenuation of obtained images that results in image sharpness decrease. For quantitative estimation of the atmosphere effect on image sharpness modulation-transfer function (MTF) of atmospheric channel may be used. In this article we analyzed earlier published experimental data for fog, haze and other phenomena, that produced changes in optical MTF of the atmosphere. We constructed the mathematical expressions and graphs that reflect the growth of attenuation of the high-frequency image components with increase of the common attenuation of an optical signal. This enables the operative MTF estimation of an atmospheric image transfer channel.*

**Key words:** optical vision, atmospheric image, transmission channel, atmosphere effect on image.

**Ключевые слова:** частотно-контрастная характеристика, передача изображения, атмосферный тракт.

**Введение**

Дистанционное телевизионное (или тепловизионное) наблюдение за объектами всегда происходит через слой атмосферы определенной протяженности, который оказывает искажающее воздействие на проходящий световой (или тепловой) поток. Это воздействие обусловлено наличием в атмосфере микро- и макрочастиц газов, жидких и твердых веществ, свойства которых изменяются под действием внешних факторов (температура, ветер, давление и т.п.). Такие взвеси частиц в воздухе приводят к образованию дымов, тумана, дымки и других помеховых явлений.

В общем, атмосфера оказывает двойное действие на оптический сигнал: во-первых, вызывает общее его ослабление, и, во-вторых, приводит к потере четкости изображения (к относительному ослаблению высокочастотных деталей) [1, 2].

Для анализа и учета искажений, возникающих в изображении, наблюдаемом через слой атмосферы определенной протяженности, удобнее всего воспользоваться «радиотехническим» подходом, когда изображение и тракт его передачи описываются двумерными функциями пространственных координат и соответственно спектральными и частотно-контрастными характеристиками изображения и элементов тракта его передачи [3-5].

Измеримой величиной в оптическом диапазоне волн обычно является средний квадрат абсолютного значения сигнала. Поэтому в задачах, подобной рассматриваемым, под оптическим сигналом во многих случаях по-

*Качество изображений, формируемых системами наблюдения (регистрации) оптического (УФ, видимого и ИК) диапазона, зависит не только от характеристик самих этих систем, но и от характеристик атмосферного тракта передачи изображения. Влияние атмосферного тракта приводит не только к общему затуханию оптического сигнала и появлению подсветки из-за рассеивания света, но и к ослаблению высокочастотных составляющих формируемого изображения, т.е. к снижению его четкости. Количественно свойства атмосферы по воздействию на четкость изображения могут быть оценены ее частотно-контрастной характеристикой (ЧКХ). В статье анализируются известные из литературы экспериментальные данные тумана, дымки и других явлений, вызывающих изменение оптической ЧКХ атмосферы. Получены математические выражения и графики, которые учитывают ослабление высокочастотных составляющих изображения, возрастающее с увеличением общего затухания оптического сигнала. Эти выражения и графики позволяют оперативно оценивать ЧКХ атмосферного тракта передачи изображения.*

нимается именно средний квадрат или интенсивность сигнала, причем выход оптических приемников выражается напряжением, пропорциональным входной интенсивности оптического сигнала.

**Эквивалентная схема оптического тракта**

Эквивалентная схема тракта передачи оптического сигнала в общем случае может быть представлена, как показано на рис. 1.

Входное ИК-изображение рассматривается как функция интенсивности, зависящая от двух пространственных переменных  $x, y$ . Энергетическое ослабление за счет действия толщи атмосферы между объектом и приемником выражается экспоненциальной зависимостью, известной как закон Бугера, согласно которому яркость  $I$  после прохождения рассеивающей среды составляет

$$I = I_0 \exp(-\varepsilon l), \quad (1)$$

где  $I_0$  – исходная яркость в изображении объекта,  $\varepsilon$  – коэффициент ослабления,  $l$  – толщина (протяженность) среды.

Таким образом, коэффициент ослабления изображения за счет атмосферы составляет

$$k = I / I_0 = \exp(-\varepsilon l), \quad (2)$$

или, в децибелах

$$k_{об} = -10 \varepsilon l g e = M \varepsilon l, \quad (2 а)$$

где  $M = l g e = 0,434294$ .

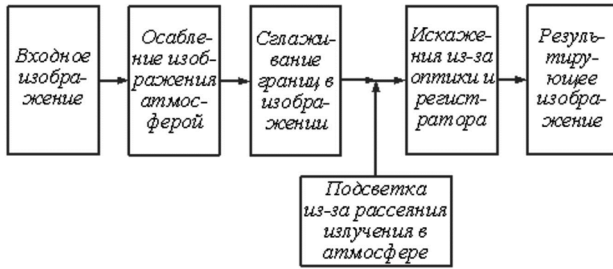


Рис. 1. Общая эквивалентная схема оптического тракта

Сглаживающее воздействие атмосферы на границы (контуры) объектов в изображении удобно описывать с помощью эквивалентных двумерных фильтров пространственных частот с ограниченной полосой пропускания.

В области реальных координат операция фильтрации выражается сверткой функции изображения с импульсной характеристикой фильтра:

$$F_1(x, y) = F(x, y) \cdot G(x, y) = \iint G(x-u, y-v) \cdot F(u, v) \cdot du \cdot dv, \quad (3)$$

где  $F(x, y)$  – исходная функция изображения;  $F_1(x, y)$  – выходная функция изображения;  $G(x, y)$  – двумерная импульсная характеристика эквивалентного фильтра.

В оптике в качестве импульсной характеристики фильтров используется функция рассеяния точки (ФРТ), либо функция рассеяния линии (ФРЛ). При экспериментальных исследованиях для получения ФРТ и ФРЛ применяют точечные и линейные источники света, которые являются техническими реализациями соответственно оптической  $\delta$ -функции и  $\delta$ -линии.

### Теоретические характеристики рассеивающих сред

Практически важным фактором, воздействующим как на видимое, так и на инфракрасное изображение, является атмосферный тракт, который содержит микро- и макрочастицы газов, жидких и твердых веществ, свойства которых изменяются под действием внешних факторов (температура, ветер, давление и т.п.). Присутствующие в воздухе взвеси частиц приводят к образованию дымов, тумана, дымки и других помеховых явлений, которые, в первую очередь, приводят к ослаблению излучений на пути от снимаемых объектов к регистратору (ИК-видеокамере). Но, как явствует из дискретной структуры указанных мешающих факторов, они должны оказывать свое воздействие и на частотно-контрастные характеристики атмосферных оптических трактов [1, 2].

Определим вначале функцию рассеяния точки (ФРТ) для самосветящегося объекта, излучение от которого попадает на приемную систему после прохождения однородного слоя дисперсной среды.

При этом ограничимся случаем больших рассеивающих частиц, когда параметр  $\rho = 2\pi a/\lambda \gg 1$ , где  $a$  – радиус частицы,  $\lambda$  – длина волны. Это выполняется для среднеквадратических диаметров частиц туманов, облаков и дождей в ультрафиолетовой, видимой и инфракрасной частях спектра (до длин волн  $\lambda = 15$  мкм).

Как показано в [1, формула 7.25], для определенного класса условий можно получить выражение ФРТ в случае прохождения излучения через однородный слой дисперсной (и кругосимметричной) среды.

В конечном итоге ФРТ получается в виде

$$D(z) = \frac{I_0 e^{-\tau}}{L^2} \cdot \frac{\rho}{2\pi} \cdot \left[ \frac{\delta(z/\rho)}{z} + \frac{\pi\rho}{4} \cdot \tau \cdot s \cdot \Lambda \cdot \frac{F(z)}{z} \right], \quad (4)$$

где  $I_0$  – сила излучения объекта,  $L$  – удаление объекта,  $\tau = \varepsilon L$ ,  $\varepsilon$  – объемный коэффициент ослабления,  $\tau S = \sigma L$ ,  $\Lambda = \sigma/\varepsilon$ ,  $\sigma$  – объемный коэффициент рассеяния,  $\delta(z/\rho)$  – дельта-функция,  $\rho = 2\pi a/\lambda$ ,

$$F(z) = \left\{ 1 - \frac{\pi z}{2} \cdot [J_0^2(z) + J_1^2(z)] + \frac{\pi}{2} \cdot J_0(z) J_1(z) \frac{\pi}{4z} + J_1^2(z) \right\}. \quad (5)$$

Оптическая передаточная функция (ОПФ) находится с помощью преобразования Фурье от ФРТ. Поскольку ФРТ имеет круговую симметрию, можно воспользоваться преобразованием Ганкеля. После введения переменной интегрирования  $z = \rho r/F$  (где  $F$  – фокусное расстояние приемной системы,  $r$  – радиус-вектор полярных координат), ОПФ получается в виде

$$M_1(x) = \frac{2\pi F^2}{\rho^2} \int_0^{\rho/x} D(z) J_0\left(\frac{x F}{\rho}\right) z dz. \quad (6)$$

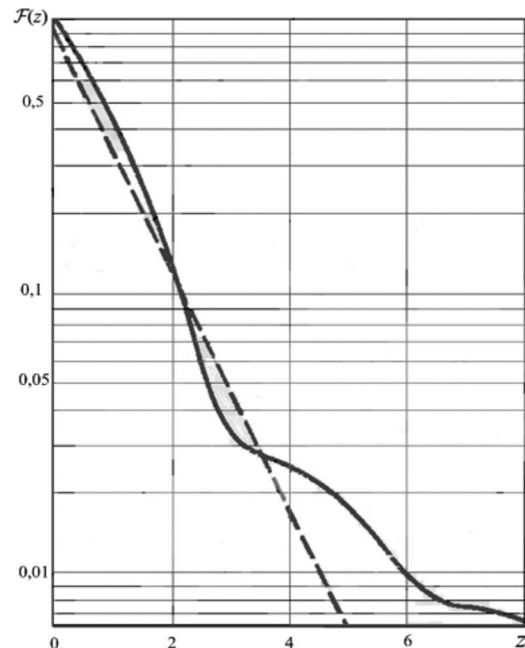


Рис. 2. График  $F(z)$  – переменной части ФРТ (сплошная линия) и его аппроксимация (штриховая прямая) [1]

График переменной части ФРТ, обозначенной  $F(z)$  приведен на рис. 2 (сплошная линия) в полулогарифмическом масштабе.

Подстановка выражения  $D(z)$  преобразует  $M_1(x)$  к двум слагаемым, первое из которых равно  $\rho$  и, следовательно, не зависит от пространственной частоты  $x$ .

Используя введенную экспоненциальную аппроксимацию  $F(z)$ , можно получить (с точностью до множителя, не зависящего от частоты) следующее выражение для переменной части ОПФ:

$$M_1(x) = \int e^{-z} J_0\left(\frac{xF}{\rho} z\right) dz = \frac{1}{\sqrt{b^2 + (xF/\rho)^2}} \quad (7)$$

В результате нормированная ОПФ  $M_1(x)/M_1(0)$  выражается представленной на рис. 3 зависимостью

$$M_1(x) / M_1(0) = (1 + k^2 x^2)^{-1/2}, \quad (8)$$

где  $k = F/\rho$ ,  $x$  – пространственная частота (при параметре  $b = 1$ ).

На рис. 3 показаны также точки, найденные путем численного интегрирования исходной функции  $F(z)$ , которые достаточно хорошо укладываются на аппроксимирующую кривую.

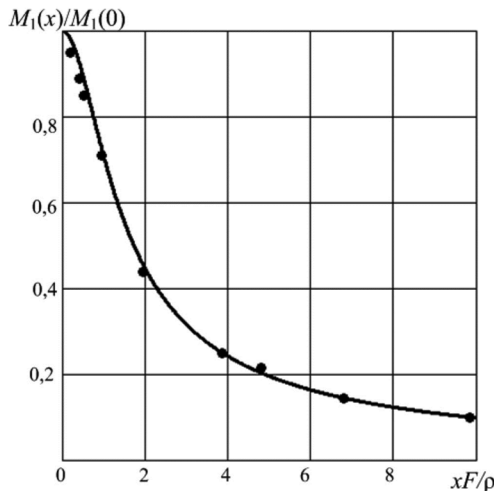


Рис. 3. График теоретической ОПФ дисперсной среды, аппроксимируемый функцией  $[1 + (xF/\rho)^2]^{-1/2}$

Таким образом, в [1] выведена общая теоретическая ОПФ рассеивающей среды, описывающая действие тумана, дымки и других атмосферных факторов.

**Экспериментальное определение ЧКХ рассеивающих сред**

Физические эксперименты по исследованию ОПФ (или ЧКХ – частотно-контрастной характеристики) рассеивающих сред описаны в работах [6] и [7].

Вначале рассмотрим результаты работы [6].

Для проведения экспериментов использовался ангар 35×20×35 м, в котором туман создавался путем испарения воды со скоростью до 1,2 т/час. В качестве тест-объекта использовалась таблица в виде чередующихся черных и белых полос, показанная на рис. 4 (размеры приведены в мм).

Самосветящийся тест-объект был выполнен в виде прозрачной панели с нанесенными на нее черными площадками, которая подсвечивалась сзади лампами

дневного света. Наблюдение (зрительное) велось с расстояния  $l = 17$  м. Частота элементов тест-объекта выражалась в периодах на градус.

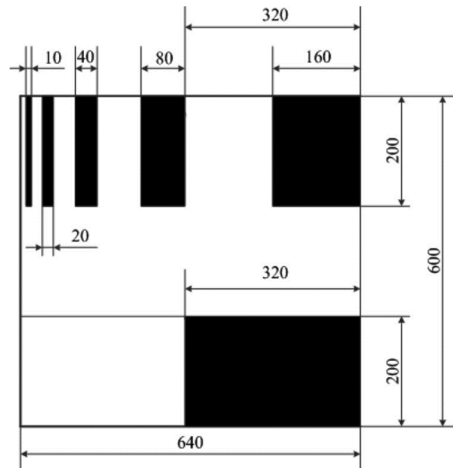


Рис. 4. Тест-объект для измерения ЧКХ тумана [6]

Численные результаты экспериментов, восстановленные из графиков в [6], приводятся в табл. 1.

Таблица 1. Результаты экспериментов по определению ЧКХ тумана [6]

$f$ , период/градус	0	0,23	0,4	0,92	2,0	Общий коэфф. передачи $M_0$
Эксперим. № 1	1,0	0,9	0,82	0,7	0,48	$4 \cdot 10^{-1}$
Эксперим. № 2	1,0	0,77	0,63	0,36		$1 \cdot 10^{-1}$
Эксперим. № 3	1,0	0,67	0,37	0,19		$6,7 \cdot 10^{-2}$
Эксперим. № 4	1,0	0,38	0,14	0,06		$1,6 \cdot 10^{-2}$

В графической форме эти результаты представлены на рис. 5. Значения ЧКХ для четырех экспериментов с разной плотностью тумана показаны значками разной формы.

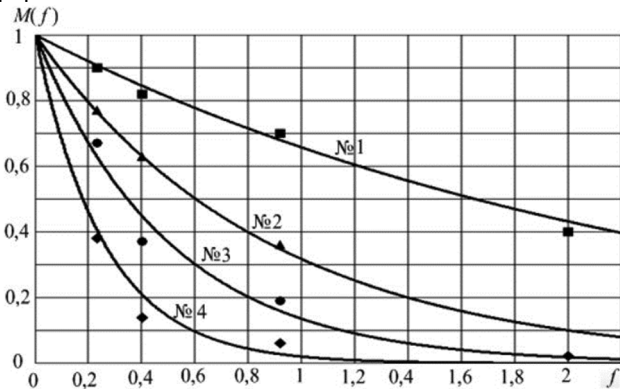


Рис. 5. Частотно-контрастные характеристики тумана; частота измеряется в периодах на градус

На рис. 5 показаны также графики экспоненциальных функций  $M_1(f) = \exp(-\alpha f)$ , которые аппроксимируют экспериментальные данные:

$$M_1(f) = \exp(-0,42 f); \quad M_2(f) = \exp(-1,15 f); \\ M_3(f) = \exp(-2,0 f); \quad M_4(f) = \exp(-3,9 f).$$

Некоторые дополнительные сведения о свойствах тумана можно получить из сопоставления общих коэффициентов передачи  $M_0$  для четырех экспериментов и

коэффициентов  $\alpha$  в показателях степени аппроксимирующих функций (табл. 2).

Таблица 2. Сопоставление затухания в тумане и крутизны спада ЧКХ

Эксперимент №	Общий коэффициент передачи тумана $M_0$	Коэффициент $\alpha$ для экспонент, аппроксимирующих ЧКХ
1	0,4	0,42
2	0,1	1,15
3	0,067	2,0
4	0,016	3,9

Для наглядности данные из табл. 2 представлены точками на рис. 6.

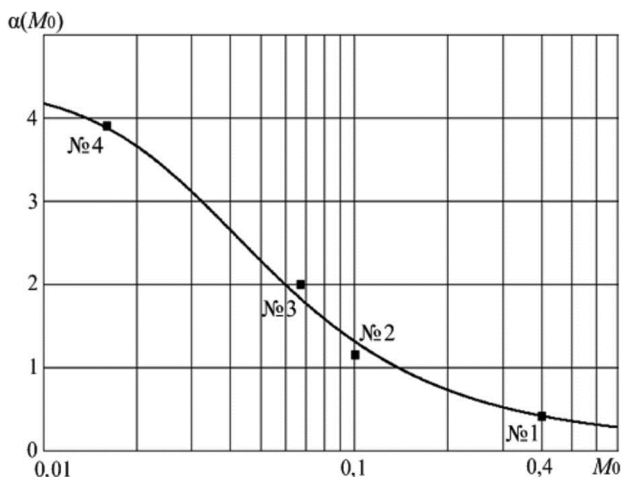


Рис. 6. Зависимость крутизны спада ЧКХ  $\alpha$  от общего коэффициента передачи  $M_0$

Для измеренных значений подобрана аппроксимирующая функция  $\alpha(M_0)$ :

$$\alpha(M_0) = \frac{4,3}{\sqrt{1 + (34M_0)^2}} + 0,1, \quad (9)$$

которая показана здесь в виде графика, иллюстрирующего зависимость скорости спада ЧКХ (выражаемой параметром  $\alpha$ ) от общего коэффициента передачи оптического сигнала  $M_0$ , т.е. фактически от протяженности рассеивающей среды. Формула (9) позволяет оценивать ЧКХ рассеивающих сред в диапазоне коэффициентов передачи оптического сигнала от 0,01 до 0,7.

Теперь рассмотрим результаты работы [7]. Здесь для экспериментального исследования ЧКХ использовались штриховые черно-белые миры, сечения которых показаны на рис. 7 в виде меандров.

Параметры графиков на рис. 7 следующие. Значение

максимальной яркости миры  $\lg B = 2,0$ . «Глубина погружения» в среде с коэффициентом ослабления  $\varepsilon = 1,0$  выражается в относительных единицах – в долях от «предельной оптической глубины обнаружения объекта  $z_0$ », которые составляют 0,2  $z_0$ , 0,4  $z_0$ , 0,6  $z_0$ , 0,8  $z_0$  и 1,0  $z_0$ .

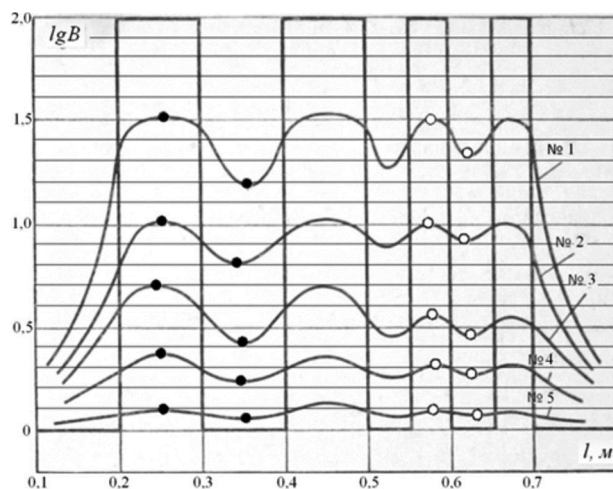


Рис. 7. Графики изменения наблюдаемой яркости штриховой миры для пространственных периодов  $l = 0,2$  и  $0,1$  м (т.е. для частот 5 и 10 период/м) [7]

При оценке ЧКХ рассеивающей среды принимается следующее.

Общий коэффициент передачи оптического сигнала  $M_0$  равен отношению наблюдаемой максимальной яркости миры  $B_{\max}$  к максимальной исходной яркости миры, равной  $B_{\text{исх}} = 100$  условных единиц.

Коэффициент передачи  $M(f)$  для переменной составляющей яркости наблюдаемой миры выражается отношением размаха переменной составляющей, т.е. разности максимальной и минимальной яркостей наблюдаемой миры ( $B_{\max} - B_{\min}$ ) к  $B_{\max}$ .

Яркость постоянной подсветки  $B_{\text{подсвет}}$  вследствие рассеяния света соответствует минимальной яркости наблюдаемой миры  $B_{\min}$  (которая в отсутствие подсветки практически была бы равна «нулевой» яркости).

На рис. 7 кружками отмечены значения яркостей  $B_{\max}$  и  $B_{\min}$ . Черные кружки соответствуют максимумам и минимумам при частоте миры 5 пер./м (периодов на метр), а белые кружки – при частоте 10 пер./м.

В табл. 3 приведены значения яркости и коэффициентов передачи для пяти экспериментов и двух значений пространственной частоты миры – 5 и 10 пер./м, полученные из графиков в [7].

Таблица 3. Данные ЧКХ среды, полученные из графиков рис. 7

График №	Яркости (вычисленные по $\lg B$ на графиках рис.7)				Коэффициенты передачи			$B_{\text{подсвет}}$
	При частоте, пер./м						Общий	
	5		10		5			
	$B_{\max}$	$B_{\min}$	$B_{\max}$	$B_{\min}$	$M(f)$			
1	33,11	15,13	31,62	21,38	0,543	0,324	0,331	15,13
2	10,47	5,13	10,47	8,32	0,510	0,205	0,105	5,13
3	4,79	2,66	3,63	2,82	0,444	0,223	0,048	2,66
4	2,25	1,70	2,09	1,86	0,244	0,110	0,022	1,70
5	1,23	1,12	1,22	1,13	0,089	0,073	0,012	1,12

Принимается, что при частоте мира  $0 \text{ пер./м}$  (постоянная составляющая) опорное значение уровня яркости мира составляет  $\lg B = 2$ , что соответствует  $10^{1.8B} = 100$  условным единицам яркости.

ЧКХ рассеивающей среды  $M(f)$  для пяти экспериментов иллюстрирует рис. 8 (частота  $f$  указана в пер./м).

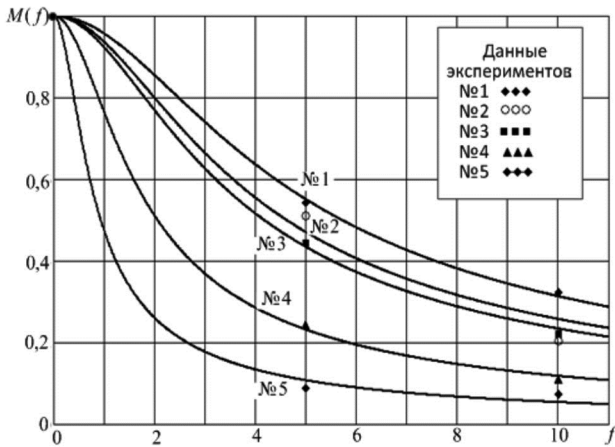


Рис. 8. ЧКХ рассеивающей среды  $M(f)$  по данным табл. 3

Представленные графические данные могут быть аппроксимированы функцией

$$M(f) = (1 + a^2 f^2)^{-1/2}, \tag{10}$$

соответствующую теоретической формуле (8), приведенной в разделе

Для показанных аппроксимирующих функций № 1-5 значения параметра  $a$  составляют:

$$a = 0,303; 0,375; 0,415; 0,84; 1,85.$$

Эти графики обеспечивают приемлемую аппроксимацию, за исключением графика № 2, что объясняется особенностями экспериментальных данных, представленных на рис. 7.

Из полученных данных [7] следует, что крутизна спада ЧКХ монотонно увеличивается с уменьшением общего коэффициента передачи  $M_0$ , как было установлено также по результатам работы [6].

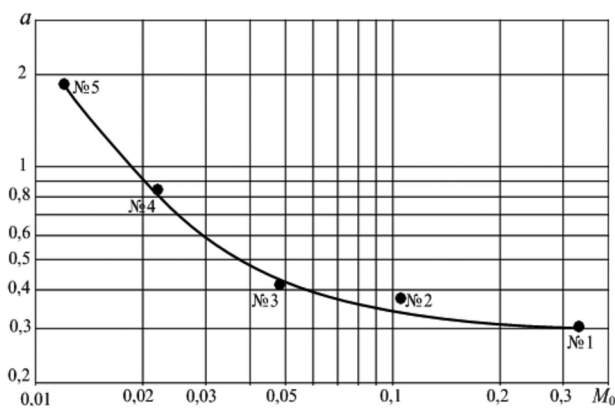


Рис. 9. Зависимость крутизны спада ЧКХ от  $M_0$

На рис. 9 представлена соответствующая эмпирическая зависимость, которая позволяет оперативно оценивать ЧКХ по известному общему коэффициенту передачи  $M_0$ .

В последних столбцах табл. 3 приводятся сведения об общем коэффициенте передачи  $M_0$  и яркости постоянной подсветки  $B_{\text{подсвет}}$ , возникающей вследствие рас-

сеивания излучения в атмосфере. В графическом виде эти сведения представлены на рис. 10.

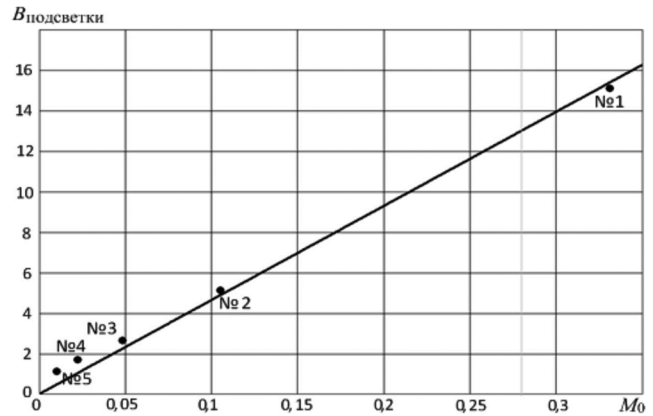


Рис. 10. Зависимость постоянной подсветки от  $M_0$

Здесь прослеживается эмпирическая линейная зависимость  $B_{\text{подсветки}}$  от  $M_0$ , выражаемая формулой  $B_{\text{подсветки}} = 41 M_0$ , которая экспериментально подтверждена до значения общего коэффициента передачи оптического сигнала  $M_0 = 0,33$ .

### Заключение

Туман, дымка и другие атмосферные явления вызывают замутнение атмосферного тракта передачи изображения, в результате чего возникает:

1) энергетическое ослабление интенсивности изображения, выражаемое общим коэффициентом передачи оптического сигнала  $M_0$ , которое приводит к снижению отношения сигнал-шум;

2) постоянная засветка изображения, приблизительно пропорциональная общему коэффициенту передачи оптического сигнала  $M_0$ ; эта засветка приводит к снижению контрастности получаемого изображения;

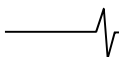
3) снижение четкости изображения, т.е. ослабление высокочастотных составляющих его спектра; это снижение описывается ЧКХ атмосферного тракта, имеющей, в общем, колоколообразный характер и зависящей от общего коэффициента передачи  $M_0$ .

Рассмотрен теоретический подход к определению зависимости ЧКХ атмосферного тракта передачи изображений от различных его характеристик. Проведен анализ известных из литературы экспериментальных характеристик тумана, дымки и других явлений, вызывающих изменение оптической ЧКХ атмосферы, и оценено их влияние на качество формируемого изображения. Экспериментальные данные в целом соответствуют исходным теоретическим положениям.

Получены аппроксимирующие математические выражения и графики, которые позволяют оперативно оценивать ЧКХ атмосферного тракта передачи изображения и учитывать их при решении практических задач моделирования и обработки изображения, наблюдаемого через протяженные слои атмосферы.

### Литература

1. Зуев В.Е. Перенос оптических сигналов в земной атмосфере (в условиях помех). / В.Е. Зуев, М.В. Кабанов – М.: Советское радио, 1977. – 368 с. – Гл. 7.



2. Зуев В.Е., Кабанов М.В. Оптика атмосферного аэрозоля. / В.Е. Зуев, М.В. Кабанов – Л.: Гидрометеоиздат, 1987.– 255 с. – Гл. 2.

3. О'Нейл Э. Введение в статистическую оптику / Э. О'Нейл – М.: Мир, 1966. – 255 с.

4. Папулис А. Теория систем и преобразований в оптике /А. Папулис. – М.: Мир, 1971. – 495 с. – Гл. 5.

5. Левшин В.Л. Пространственная фильтрация в оптических системах пеленгации / В.Л. Левшин – М.:

Советское радио, 1971. – 200 с. – Гл. 1.

6. Naoyoshi Nameda. Fog modulation transfer function and signal lighting / Nameda Naoyoshi – Lighting Research and Technology. – 1992. – Vol. 24, № 2. – Pp. 103-106.

7. Смирнов В.А. Теория и метод решения задач о переносе оптического изображения в рассеянных средах / В.А. Смирнов – Вопросы радиоэлектроники, серия Техника телевидения. – 1965, вып. 6. – С. 109-124.