

# ЦИФРОВАЯ ОБРАБОТКА СИГНАЛОВ

Научно-технический журнал  
№ 3/2014

Издается с 1999 года  
Выходит четыре раза в год

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

## ЗАМЕСТИТЕЛИ ГЛАВНОГО РЕДАКТОРА:

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

## РЕДАКЦИОННАЯ КОЛЛЕГИЯ:

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

## Адрес редакции:

г. Москва, ул. Авиамоторная, д. 8  
Научный центр МТУСИ  
Тел.: 8 903 221 79 79  
E-mail: rntores@mail.ru  
vityazev.v.v@rsreu.ru  
http://www.dspa.ru

## Издатель:

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

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

Подписано в печать 22.09.14 г.  
Формат 60x90/8.

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

Заказ № 6147. Отпечатано в  
ООО НПЦ «Информационные технологии»  
Рязань, ул. Островского, д. 21/1  
тел.: (4912) 98-69-84

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

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

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

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

### В НОМЕРЕ:

#### Самойлин Е.А., Шипко В.В.

Метод различения случайных сигналов  
многокомпонентных изображений и импульсных помех  
на основе свойства межканальной избыточности ..... 2

#### Дрынкин В.Н., Царева Т.И.

Метод повышения разрешающей способности изображений .... 9

#### Новиков А.И., Ефимов А.И.

Предварительное совмещение изображений  
и методы оценки качества совмещения ..... 15

#### Матюнин С.Б., Ватолин Д.С.

Обработка областей наложенных объектов  
при построении карт глубины для видео ..... 23

#### Демидова Л.А., Тишкин Р.В., Труханов С.В.

Алгоритмы идентификации гиперспектральных  
характеристик объектов в задачах  
дистанционного зондирования Земли ..... 30

#### Еремеев В.В., Макаренков А.А., Москвитин А.Э.

Повышение пространственного разрешения материалов  
гиперспектральной съемки Земли на основе  
их комплексирования с высокодетальными снимками ..... 38

#### Алпатов Б.А., Бабаян П.В., Смирнов С.А., Масленников Е.А.

Алгоритм предварительного оценивания  
пространственной ориентации объекта  
с помощью дескриптора внешнего контура ..... 43

#### Медведева Е.В., Карлушин К.А.

Метод выделения движущихся объектов в видеопотоке  
и оценка точности определения их координат ..... 47

#### Муравьев В.С., Фельдман А.Б.

Повышение эффективности селекции объектов  
в системах видеослежения при наличии яркостных помех ..... 51

#### Холопов И.С.

Алгоритм упрощенной компенсации дисторсии  
при проецировании видеоизображения на асферические  
отражающие поверхности априорно неизвестной формы ..... 57

#### Медведев Д.Г.

Метод выделения контуров объектов  
с нечеткими краями на цветных изображениях ..... 62

#### Акинин М.В.

Нейросетевой алгоритм уточнения векторных  
топографических карт по данным дистанционного  
зондирования Земли ..... 68

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

## МЕТОД РАЗЛИЧЕНИЯ СЛУЧАЙНЫХ СИГНАЛОВ МНОГОКОМПОНЕНТНЫХ ИЗОБРАЖЕНИЙ И ИМПУЛЬСНЫХ ПОМЕХ НА ОСНОВЕ СВОЙСТВА МЕЖКАНАЛЬНОЙ ИЗБЫТОЧНОСТИ

*Самойлин Е.А., д.т.н., профессор ВУНЦ ВВС «Военно-воздушная академия им. проф. Н.Е. Жуковского и Ю.А. Гагарина», г. Воронеж, e-mail: es977@mail.ru*

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

**Ключевые слова:** многокомпонентные изображения, импульсные помехи, межканальная избыточность, случайные сигналы, градиентная оценка.

### Введение

Для многих современных систем и приложений характерно использование многокомпонентных цифровых изображений, к которым можно отнести цветные оптические, многочастотные и многополяризационные радиолокационные, мульти- и гиперспектральные изображения. Многокомпонентность таких данных позволяет либо повысить точность и надежность решения ряда конечных задач, либо извлекать полезную информацию, которую принципиально невозможно извлечь из одноканальных изображений [1].

Для многих из перечисленных выше типов многокомпонентных изображений одним из основных факторов, снижающих качество исходных данных, является наличие различного рода помех и искажающих факторов. Поэтому для повышения визуального качества изображений и, соответственно, эффективности решения конечных задач (классификации, обнаружения объектов, измерения параметров и т.п.) используют алгоритмы фильтрации изображений, синтезируемые для конкретной модели помех [2].

Как известно [3], многие практические условия передачи и преобразования цифровых изображений становятся причиной появления на растре случайных помех, относящихся к классу точечных или импульсных. При этом импульсные помехи (ИП) со случайными значениями импульсов являются наиболее сложным случаем при их фильтрации [4, 5]. Наиболее распространенными на практике являются ИП с равномерным, бимодальным и модальным гауссовоподобными распределениями яркости [3-5].

Практически все известные [6-10] алгоритмы подавления ИП на изображениях на первом этапе обработки реализуют процедуру различения полезных сигналов и ИП, а на втором – восстановление каким-либо методом потерянных (вследствие воздействия помех) элементов изображения. К настоящему времени разработано множество разнообразных алгоритмов обнаружения ИП [5, 6, 10, 11], являющихся покомпонентными. При этом лишь незначительное количество этих алгоритмов мо-

*Предложен метод различения случайных сигналов многокомпонентных изображений и импульсных помех, основанный на вычислении и сопоставлении межканальных градиентных оценок. Представлены результаты численных исследований метода, свидетельствующие о снижении числа ошибок различения сигналов и помех по сравнению с известными ранговыми алгоритмами обработки многокомпонентных изображений.*

жет быть обобщено на векторный случай. Поэтому представляет интерес использование свойства межканальной избыточности для построения векторного алгоритма обнаружения ИП. В частности, это свойство было использовано на этапе восстановления сигналов многокомпонентных изображений [6, 7]. Однако в работах [6, 7] на первом этапе обработки использованы известные скалярные обнаружители ИП.

Цель предлагаемой работы – разработка метода различения случайных сигналов и ИП на цифровых многокомпонентных изображениях на основе свойства межканальной избыточности.

### Постановка задачи различения сигналов и помех на многокомпонентных цифровых изображениях

Используемая модель исходного оцифрованного по строкам  $i$  и столбцам  $j$   $L$ -компонентного изображения  $\Lambda$  в общем случае имеет вид [3]

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

Здесь  $i = 1, \dots, m, j = 1, \dots, n$ , где  $m, n$  – число строк и столбцов изображения соответственно;  $l = 1, \dots, L$ , где  $L$  – число компонент,  $N$  – степень квантования яркости компонент изображения.

Модель  $L$ -компонентного изображения, искаженного воздействием ИП, описывается выражением

$$\mathbf{X} = [x_{i,j}^1 \quad x_{i,j}^2 \quad \dots \quad x_{i,j}^L]^T, \quad (2)$$

где каждый канал  $l$  (2) содержит ИП

$$x_{i,j}^l = \begin{cases} h_{i,j}^l & \text{с вероятностью } p_h^l, \\ \lambda_{i,j}^l & \text{с вероятностью } 1 - p_h^l, \end{cases} \quad (3)$$

где  $p_h^l$  – вероятность появления помехи  $h_{i,j}^l$  в ячейке с координатой  $(i, j)$ , которая не зависит ни от наличия помех в других координатах, ни от исходного сигнала, ни от

наличия помех в соседних каналах;  $h_{i,j}^l$  – независимые случайные значения ИП в каналах изображения. Эти значения описываются, как правило, равномерным законом распределения и соответствуют интервалу квантования (1):

$$h_{i,j}^l = \text{md}[0, \dots, 2^N - 1]. \quad (4)$$

Необходимо синтезировать алгоритм различения сигналов и помех, позволяющий сформировать бинарные матрицы оценок положения ИП  $\hat{h}_{i,j}^l \in [0, 1]$  на изображениях  $x_{i,j}^l$ , наиболее близких к истинному положению помех  $\tilde{h}_{i,j}^l = \begin{cases} 1, & x_{i,j}^l = h_{i,j}^l, \\ 0, & x_{i,j}^l = \lambda_{i,j}^l, \end{cases}$  где единицы указывают на координаты ИП, а нули – на полезные сигналы изображения.

**Сущность метода различения сигналов и помех на основе межканальной избыточности**

**Сущность метода различения сигналов и помех на основе межканальной избыточности**

Как известно [1-3], цветные RGB-компонентные изображения являются частным случаем многокомпонентных изображений, поэтому предлагаемый метод рассмотрим для случая трехкомпонентного цифрового изображения с цветовыми каналами R (красный), G (зеленый) и B (синий).

Введем следующие обозначения. Пусть  $a = (R \ V \ G \ V \ B)$ ,  $b = (R \ V \ G \ V \ B)$ ,  $c = (R \ V \ G \ V \ B)$ , при этом  $a \neq b$ ,  $c$ ;  $b \neq a$ ,  $c$  и  $c \neq a$ ,  $b$ , символ V означает логическое «ИЛИ».

Тогда модель изображения, искаженного ИП (2), примет вид

$$\mathbf{X} = \begin{bmatrix} x_{i,j}^a & x_{i,j}^b & x_{i,j}^c \end{bmatrix}^T. \quad (5)$$

Обработка многокомпонентного изображения X осуществляется скользящим по пространству окном (апертурой) с параметрами  $s = (0, \dots, S)$  и  $q = (0, \dots, Q)$ , определяющими ее размер по строкам  $i$  и столбцам  $j$  соответственно.

На 1-м этапе обработки в каждом цветовом канале вычисляются  $(2S + 1)(2Q + 1) - 1$  локальных градиента в направлении от центрального элемента апертуры  $(i, j)$  до каждого из остальных ее элементов в пределах окрестности  $i \pm s, j \pm q$ :

$$g_{i \pm s, j \pm q}^a = x_{i,j}^a - x_{i \pm s, j \pm q}^a; \quad (6)$$

$$g_{i \pm s, j \pm q}^b = x_{i,j}^b - x_{i \pm s, j \pm q}^b; \quad (7)$$

$$g_{i \pm s, j \pm q}^c = x_{i,j}^c - x_{i \pm s, j \pm q}^c, \quad (8)$$

при этом переменные  $s, q$  одновременно не равны нулю.

После этого вычисляется набор невязок (межканальных градиентных ошибок) локальных градиентов между всеми цветовыми компонентами:

$$\Delta_{i \pm s, j \pm q}^{ab} = |g_{i \pm s, j \pm q}^a - g_{i \pm s, j \pm q}^b|; \quad (9)$$

$$\Delta_{i \pm s, j \pm q}^{bc} = |g_{i \pm s, j \pm q}^b - g_{i \pm s, j \pm q}^c|; \quad (10)$$

$$\Delta_{i \pm s, j \pm q}^{ac} = |g_{i \pm s, j \pm q}^a - g_{i \pm s, j \pm q}^c|. \quad (11)$$

На 2-м этапе обработки на основе значений межканальных градиентных ошибок (9)-(11) осуществляется определение элементов, подозрительных на помехи.

Так, в случае если взаимная межканальная градиентная ошибка каких-либо двух каналов превышает некоторое пороговое значение, то принимается предварительное решение о наличии помехи в этих каналах. С этой целью вначале формируются бинарные оценки межканальных градиентных ошибок по каждому направлению:

$$\hat{h}_{i \pm s, j \pm q}^{ab} = \begin{cases} 1, & \Delta_{i \pm s, j \pm q}^{ab} > T, \\ 0, & \Delta_{i \pm s, j \pm q}^{ab} \leq T, \end{cases} \quad (12)$$

$$\hat{h}_{i \pm s, j \pm q}^{bc} = \begin{cases} 1, & \Delta_{i \pm s, j \pm q}^{bc} > T, \\ 0, & \Delta_{i \pm s, j \pm q}^{bc} \leq T, \end{cases} \quad (13)$$

$$\hat{h}_{i \pm s, j \pm q}^{ac} = \begin{cases} 1, & \Delta_{i \pm s, j \pm q}^{ac} > T, \\ 0, & \Delta_{i \pm s, j \pm q}^{ac} \leq T, \end{cases} \quad (14)$$

где  $T$  – порог, выбираемый экспериментально, в зависимости от корреляции между компонентами изображения X.

А затем, по полученным наборам оценок (12)-(14) формируются суммарные оценки для всех возможных пар каналов каждого сигнала изображения:

$$\hat{h}_{i,j}^{ab} = \begin{cases} 1, & \left( \sum_{s=0}^S \sum_{q=0}^Q \hat{h}_{i \pm s, j \pm q}^{ab} \right) > t, \\ 0, & \left( \sum_{s=0}^S \sum_{q=0}^Q \hat{h}_{i \pm s, j \pm q}^{ab} \right) \leq t, \end{cases} \quad (15)$$

$$\hat{h}_{i,j}^{bc} = \begin{cases} 1, & \left( \sum_{s=0}^S \sum_{q=0}^Q \hat{h}_{i \pm s, j \pm q}^{bc} \right) > t, \\ 0, & \left( \sum_{s=0}^S \sum_{q=0}^Q \hat{h}_{i \pm s, j \pm q}^{bc} \right) \leq t, \end{cases} \quad (16)$$

$$\hat{h}_{i,j}^{ac} = \begin{cases} 1, & \left( \sum_{s=0}^S \sum_{q=0}^Q \hat{h}_{i \pm s, j \pm q}^{ac} \right) > t, \\ 0, & \left( \sum_{s=0}^S \sum_{q=0}^Q \hat{h}_{i \pm s, j \pm q}^{ac} \right) \leq t, \end{cases} \quad (17)$$

где  $t$  – порог, выбираемый экспериментально в зависимости от размера апертуры фильтра.

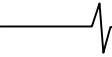
Заключительным этапом обработки многокомпонентного изображения является определение по полученным подозрительным элементам (т.е. парным оценкам (15)-(17)) конечных оценок пространственного положения ИП в цветовых компонентах. С этой целью происходит логическое сравнение парных оценок (15)-(17), которое позволяет получить конечные межканальные оценки шумовых элементов каждой компоненты в отдельности:

$$\hat{h}_{i,j}^a = \begin{cases} 1, & (\hat{h}_{i,j}^{ab} = 1) \wedge (\hat{h}_{i,j}^{ac} = 1), \\ 0, & \text{в остальных случаях,} \end{cases} \quad (18)$$

$$\hat{h}_{i,j}^b = \begin{cases} 1, & (\hat{h}_{i,j}^{ab} = 1) \wedge (\hat{h}_{i,j}^{bc} = 1), \\ 0, & \text{в остальных случаях,} \end{cases} \quad (19)$$

$$\hat{h}_{i,j}^c = \begin{cases} 1, & (\hat{h}_{i,j}^{bc} = 1) \wedge (\hat{h}_{i,j}^{ac} = 1), \\ 0, & \text{в остальных случаях,} \end{cases} \quad (20)$$

где символ  $\wedge$  означает логическое «И». Таким образом, полученные бинарные оценки (18)-



(20) указывают на наличие ( $\hat{h}_{i,j}^{a,b,c} = 1$ ) или отсутствие ( $\hat{h}_{i,j}^{a,b,c} = 0$ ) ИП в соответствующих координатах  $(i, j)$  компонент  $a, b, c$  изображения  $X$ .

### Результаты вычислительных исследований предлагаемого метода

Вычислительные исследования проводились на разнообразных цветных (RGB-компонентных) цифровых оптических изображениях со следующими параметрами  $m \times n = 320 \times 320$ ,  $N = 8$ . Приводимые далее результаты усреднены по выборке более чем из 100 самых различных цветных цифровых изображений с указанными параметрами.

Для обоснования выбора величины порога  $T$  в формулах (12)-(14) на рис. 1(а) представлены рассчитанные нормированные распределения межканальных градиентных ошибок  $F(\Delta) = \frac{\Delta_{i,j}}{mn}$ , где расчет ошибок  $\Delta_{i,j}$  выполнялся по единственному направлению от центра апертуры

$$\Delta_{i,j}^1 = \frac{1}{3} (\Delta_{i\pm s, j\pm q}^{ab} + \Delta_{i\pm s, j\pm q}^{ac} + \Delta_{i\pm s, j\pm q}^{bc}). \quad (21)$$

На рис. 1(б) представлено рассчитанное распределение  $F(\Delta)$  суммарных градиентных ошибок по всем 8-ми направлениям апертуры  $3 \times 3$ :

$$\Delta_{i,j}^8 = \frac{\sum_{s=0}^S \sum_{q=0}^Q (\Delta_{i\pm s, j\pm q}^{ab} + \Delta_{i\pm s, j\pm q}^{ac} + \Delta_{i\pm s, j\pm q}^{bc})}{3((2S+1)(2Q+1)-1)}. \quad (22)$$

На рис. 1(а, б) кривая 1 – для незашумленных компонент изображения; кривая 2 – при  $p_h^a = 0,1$  в одной из компонент; кривая 3 – при  $p_h^a = 0,5$  в одной из компонент; кривая 4 – при  $p_h^a = 0,8$  в одной из компонент.

Из рис. 1(а, б) видно, что градиентные ошибки для незашумленных изображений (кривая 1) в большинстве случаев группируются на интервале  $\Delta = 0 \dots 7$ , при этом для случая зашумления одного из каналов с ростом интенсивности помех происходит смещение центра распределения градиентных ошибок в сторону увеличения (кривые 2-4). Отсюда следует, что при правильном выборе порога в (12)-(14) возможно определить, каким именно фактором обусловлено значение градиента – наличием помехи или присутствием в апертуре перепада полезных сигналов изображения. Таким образом, величину порога  $T$  целесообразно устанавливать как максимальное отклонение градиентной ошибки от ее математического ожидания  $m_\Delta$ . Очевидно, что для незашумленных компонент изображения  $X$ ,  $m_\Delta = 0$ . При усреднении результатов множества вычислительных исследований установлено наилучшее значение порога, равное  $T = 7$ .

С целью обоснования выбора величины порога  $t$  в ходе исследований вычислялась общая частота ошибок обнаружения

$$\delta_{\text{ош}}^{a,b,c} = \delta_{\text{ПП}}^{a,b,c} + \delta_{\text{ЛТ}}^{a,b,c}, \quad (23)$$

где  $\delta_{\text{ПП}}^{a,b,c}$ ,  $\delta_{\text{ЛТ}}^{a,b,c}$  – частоты ошибок обнаружения ИП, соот-

ветствующие пропуску помех (ПП) и ложной тревоге (ЛТ):

$$\delta_{\text{ПП}}^{a,b,c} = \frac{1}{mn} \sum_{i=1}^m \sum_{j=1}^n \begin{cases} 1, & (\tilde{h}_{i,j}^{a,b,c} - \hat{h}_{i,j}^{a,b,c}) = 1, \\ 0, & (\tilde{h}_{i,j}^{a,b,c} - \hat{h}_{i,j}^{a,b,c}) \neq 1, \end{cases} \quad (24)$$

$$\delta_{\text{ЛТ}}^{a,b,c} = \frac{1}{mn} \sum_{i=1}^m \sum_{j=1}^n \begin{cases} 1, & (\hat{h}_{i,j}^{a,b,c} - \tilde{h}_{i,j}^{a,b,c}) = 1, \\ 0, & (\hat{h}_{i,j}^{a,b,c} - \tilde{h}_{i,j}^{a,b,c}) \neq 1. \end{cases} \quad (25)$$

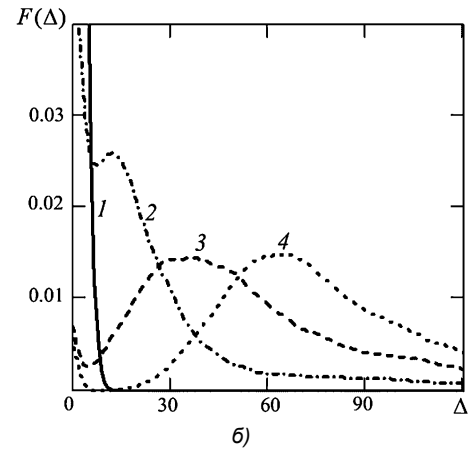
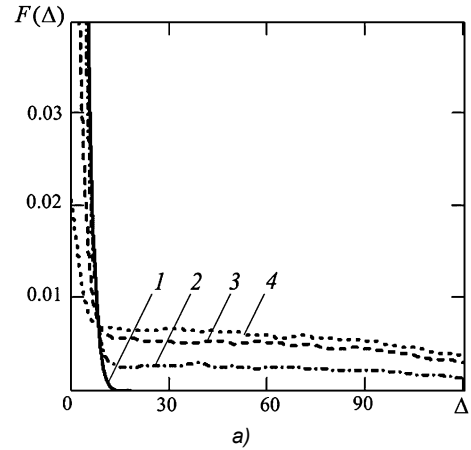


Рис. 1. Плотность распределения межканальных градиентных ошибок по одному направлению (а) и всем 8-ми направлениям (б) для апертуры  $3 \times 3$

На рис. 2 представлены зависимости значений ошибки  $\delta_{\text{ош}}^a$  от вероятности наличия ИП  $p_h^a$  с равномерным распределением яркости в канале  $a$  при условии отсутствия помех в каналах  $b$  и  $c$  ( $p_h^a = 0, \dots, 1$ ;  $p_h^b = 0$ ;  $p_h^c = 0$ ) для апертуры  $3 \times 3$ , и  $T = 7$ .

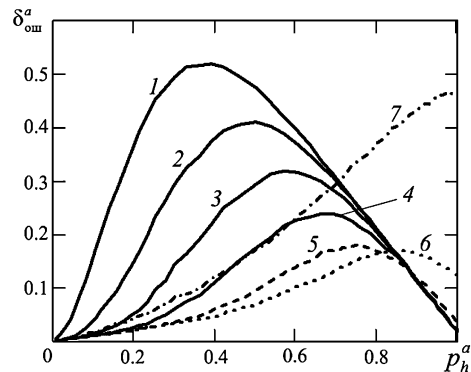


Рис. 2. Зависимость  $\delta_{\text{ош}}^a$  от  $p_h^a$ : 1 –  $t = 1$ ; 2 –  $t = 2$ ; 3 –  $t = 3$ ; 4 –  $t = 4$ ; 5 –  $t = 5$ ; 6 –  $t = 6$ ; 7 –  $t = 7$



Из рис. 2 видно, что наименьшая частота ошибок  $\delta_{\text{ош}}^a$  соответствует величине порога  $t = 6$  (кривая 6). При проведении аналогичных исследований для апертур других размеров ( $S, Q = 2 \dots 5$ ) установлена общая зависимость для значения порога, приводящего к минимальной частоте ошибок  $\delta_{\text{ош}}^{a,b,c}$ :

$$t = (2S + 1)(2Q + 1) - SQ - 2. \quad (26)$$

Для исследования качественных характеристик предлагаемого метода вычислялись ошибки типа ПП и ЛТ различных обнаружителей в соответствии с выражениями

$$\delta_{\text{ПП}} = \frac{1}{3} (\delta_{\text{ПП}}^a + \delta_{\text{ПП}}^b + \delta_{\text{ПП}}^c); \quad (27)$$

$$\delta_{\text{ЛТ}} = \frac{1}{3} (\delta_{\text{ЛТ}}^a + \delta_{\text{ЛТ}}^b + \delta_{\text{ЛТ}}^c). \quad (28)$$

Введем параметр  $p = 0, \dots, 1$ , в соответствии с которым изменяются заданные вероятности наличия ИП ( $p_h^a, p_h^b, p_h^c$ ).

На рис. 3-5 приведены ошибки обнаружения (27), (28) ИП с равномерным распределением яркости на интервале (0...255) для алгоритма на основе предлагаемого метода, а также для известных ранговых процедур. На рис. 3-5 кривая 1 – известный разностный алгоритм, основанный на выделении границ объектов [11], где решение о принадлежности элемента  $x_{i,j}^a$  к контуру принимается, исходя из разности между значением яркости этого элемента и оценкой среднего значения апертуры; кривая 2 – известный алгоритм, основанный на сравнении центрального элемента апертуры  $x_{i,j}^a$  с медианным элементом [11]; кривая 3 – предлагаемый алгоритм с апертурой 3x3; кривая 4 – предлагаемый алгоритм с апертурой 5x5.

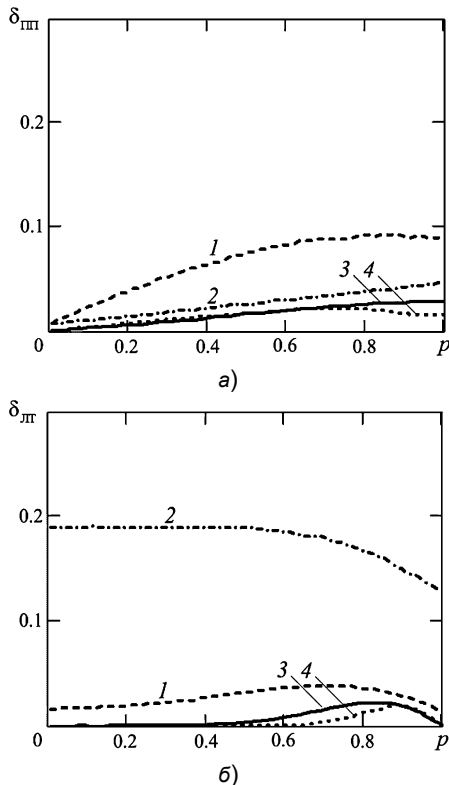


Рис. 3. Зависимость ошибок  $\delta_{\text{ПП}}$  (а) и  $\delta_{\text{ЛТ}}$  (б) от  $p$  при наличии ИП в одном канале ( $p_h^a = 0, \dots, 1; p_h^b = 0; p_h^c = 0$ )

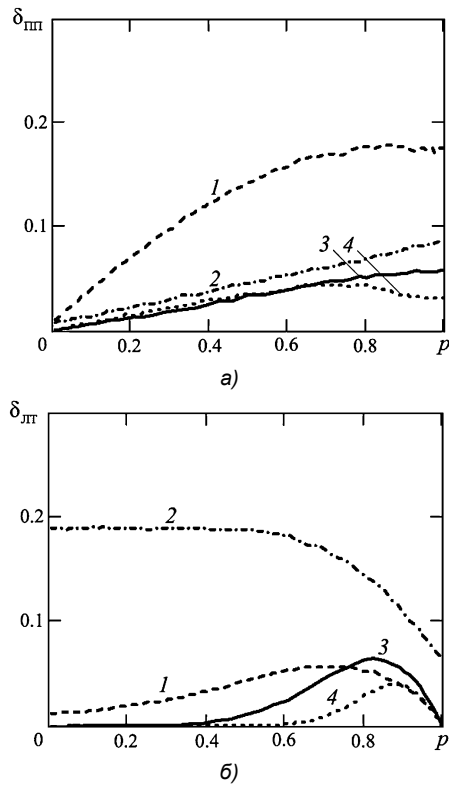


Рис. 4. Зависимость ошибок  $\delta_{\text{ПП}}$  (а) и  $\delta_{\text{ЛТ}}$  (б) от  $p$  при наличии ИП в двух каналах ( $p_h^a = 0, \dots, 1; p_h^b = 0, \dots, 1; p_h^c = 0$ )

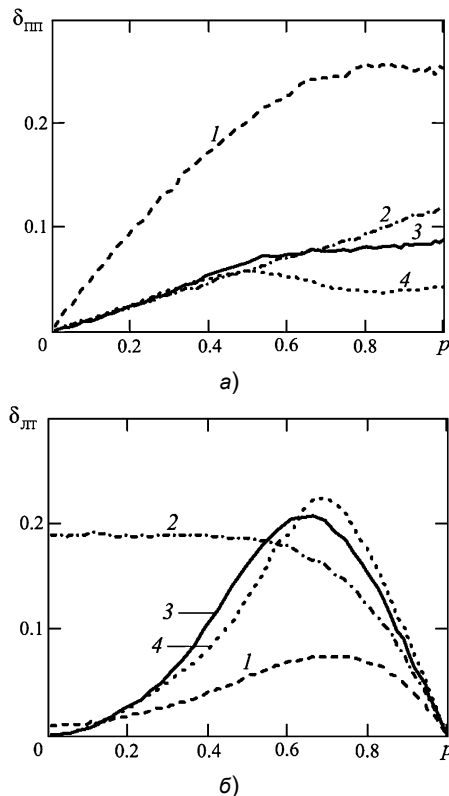
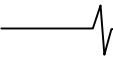


Рис. 5. Зависимость ошибок  $\delta_{\text{ПП}}$  (а) и  $\delta_{\text{ЛТ}}$  (б) от  $p$  при наличии ИП в каждом канале ( $p_h^a = 0, \dots, 1; p_h^b = 0, \dots, 1; p_h^c = 0, \dots, 1$ )

Полученные в результате вычислительных исследований значения ошибок обнаружения  $\delta_{\text{ПП}}$  и  $\delta_{\text{ЛТ}}$  разработанного межканального метода (кривые 3, 4) более низкие по сравнению с известными покомпонентными алго-



ритмами (кривые 1, 2), за исключением величин ошибок  $\delta_{ЛТ}$  для случая, когда ИП присутствуют во всех каналах рис. 5(б). Как представляется, рост ошибок  $\delta_{ЛТ}$  (рис. 5, кривые 3, 4) возникает из-за наличия ИП, которые являются коррелированными между всеми тремя компонентами. При этом увеличение размера апертуры (кривая 4) заметно снижает значения ошибок ПП и ЛТ, так как возрастает число элементов, по которым рассчитываются градиенты.

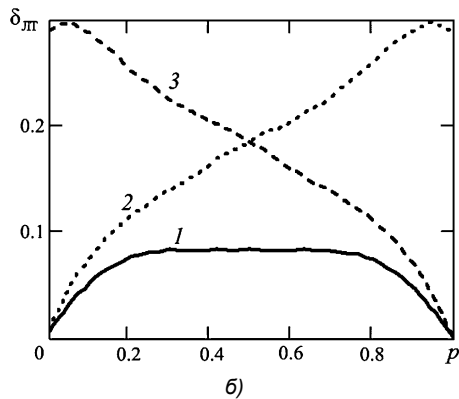
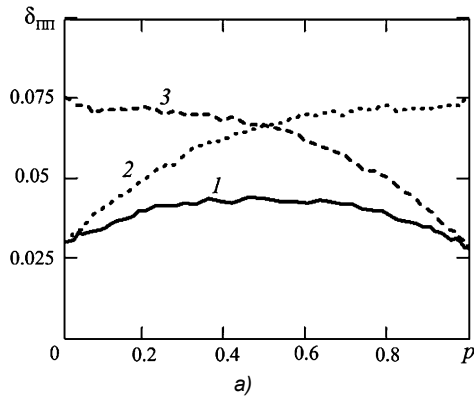


Рис. 6. Зависимость ошибок  $\delta_{ПП}$  (а) и  $\delta_{ЛТ}$  (б) от  $p$ , в соответствии с изменением вероятностей наличия ИП для случаев ( $p_h^a = 0, \dots, 1; p_h^b = 1, \dots, 0; p_h^c = 0$ ) – 1, ( $p_h^a = 0, \dots, 1; p_h^b = 0, \dots, 1; p_h^c = 1, \dots, 0$ ) – 2, ( $p_h^a = 0, \dots, 1; p_h^b = 1, \dots, 0; p_h^c = 1, \dots, 0$ ) – 3

Из рис. 3-5 следует, что предлагаемый метод позволяет получить наименьшую ошибку обнаружения как для случая зашумления одного из каналов, так и наличия ИП во всех цветовых компонентах цифрового изображения.

С целью более объективного анализа разработанного метода на рис. 6 приведены ошибки обнаружения (27), (28) для еще некоторых случаев изменения вероятностей наличия ИП с равномерным распределением яркости в цветовых каналах изображения. Ошибки обнаружения (кривые 1-3) на рис. 6 характеризуются достаточно низким уровнем для различных вероятностей наличия ИП во всех каналах изображения.

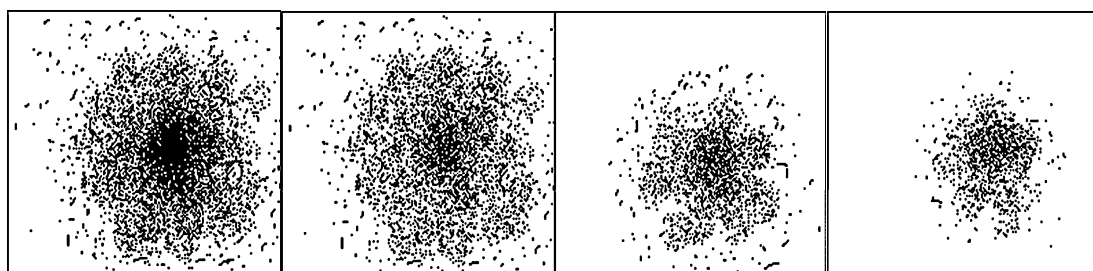
На рис. 7 показаны результаты обнаружения ИП с гауссовским распределением по пространству и усеченным гауссовским верхнемодальным распределением яркости с модой 255 и дисперсией 20 в каждом канале цветного изображения. Такая модель помех на изображениях соответствует лазерной засветке оптико-электронной системы с RGB каналами. На рис. 7 а – цветное изображение и его R, G и B компоненты с помехами, б – соответствующие бинарные изображения истинного положения ИП  $\tilde{h}_{i,j}^{a,b,c}$  на изображениях, в – бинарные изображения

обнаруженных ИП  $\hat{h}_{i,j}^{a,b,c}$  разработанным межканальным алгоритмом с апертурой  $3 \times 3$ , г – бинарные изображения обнаруженных ИП  $\hat{h}_{i,j}^{a,b,c}$  разработанным межканальным алгоритмом с апертурой  $5 \times 5$ .

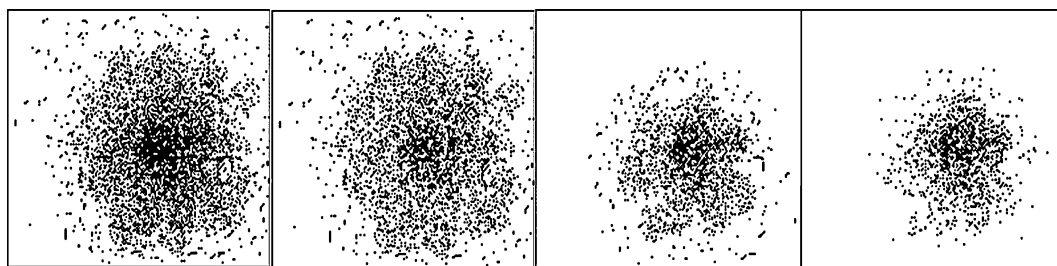
Как известно [3], ИП являются частным случаем аппликативных помех [12], действие которых проявляется в затенении (замещении) отдельных участков полезного изображения фрагментами со случайной амплитудой, текстурой и формой. Аппликативные помехи связаны с появлением различного рода неоднородностей, локальных областей аномальных значений, областей закрытия и пораженных участков. На рис. 8 приведены результаты обнаружения пространственно коррелированных аппликативных искажений с усеченным гауссовским нижнемодальным распределением яркости с модой 0 и дисперсией 20 в каждом канале цветного изображения. Буквенные обозначения на рис. 8 соответствуют обозначениям на рис. 7.



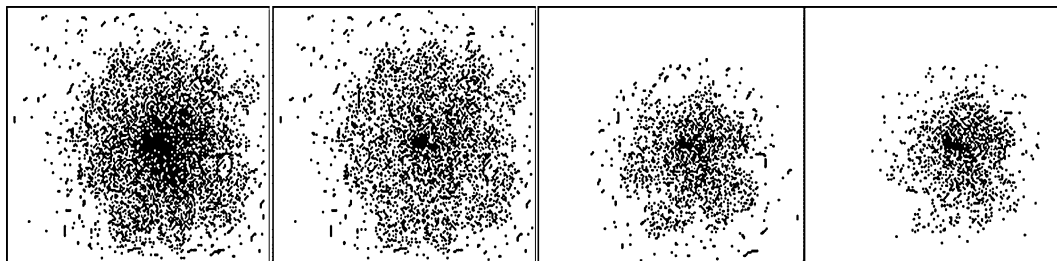
а)



б)



в)



а)

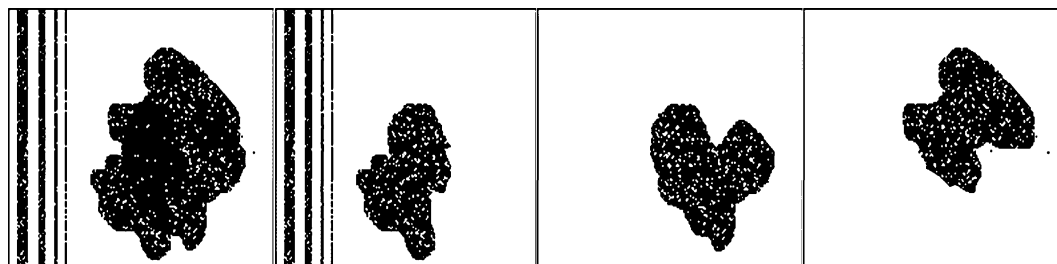
Рис. 7. Результаты обнаружения ИП на цветных изображениях



а)



б)

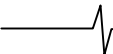


в)



а)

Рис. 8. Результаты обнаружения пространственно коррелированных аппликативных искажений на цветных изображениях



## Заключение

Как показали результаты вычислительных исследований, свойство межканальной избыточности многокомпонентных цифровых изображений может быть использовано в задачах различения полезных сигналов и ИП, решаемых на первом этапе при цифровой обработке и анализе изображений. Разработанный метод является достаточно универсальным к таким параметрам ИП как закон распределения яркости и вероятность их наличия. Кроме того, разработанный метод применим для выделения протяженных аномалий или аппликативных искажений, а также в следящих системах для обнаружения и сопровождения точечных объектов.

## Литература

1. Lukin V. // Proc. of NATO Advanced Research Workshop on Geographical Information Processing of Multichannel RS data for Environment Monitoring // Proc. of NATO Advanced Research Workshop on Geographical Information Processing and Visual Analytics for Environmental Security. Trento, (Italy): Springer Netherlands, July 2009. P. 129–138.
2. Белоусов Ю.И., Фисенко В.Т. Особенности алгоритмов цифровой обработки изображений в оптико-электронных системах нового поколения // Изв. вузов. Приборостроение. 2009. Т. 52. № 8. С. 7-12.
3. Гонсалес Р., Вудс Р. Цифровая обработка изображений. – М.: Техносфера, 2005. – 1072 с.
4. Хуанг Т.С., Эклунд Дж.-О., Нуссбаумер Г. Дж., Зохар Ш., Юстуссон Б.И., Тянь Ш.-Г. Быстрые алгоритмы в цифровой обработке изображений. – М.: Радио и связь, 1984. – 224 с.
5. Самойлин Е.А. Оптимальное оценивание положения негауссовских импульсных помех на изображениях // Автометрия. 2009. Т. 45. № 3. С. 74-82.
6. Самойлин Е.А., Шипко В.В. Метод межканальной компенсации импульсных помех в задачах восстановления многокомпонентных цифровых изображений // Оптический журн. 2013. Т. 80. № 10. С. 53.
7. Самойлин Е.А., Шипко В.В. Итерационные алгоритмы межканальной градиентной реконструкции мно-

гокомпонентных изображений, искаженных аппликативными помехами // Оптический журнал. 2014. Т. 81. № 4. С. 54-60.

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

9. Воскобойников Ю.Е., Белявцев В.Г. Нелинейные алгоритмы фильтрации векторных сигналов // Автометрия. 1999. № 5. С. 97–105.

10. Самойлин Е.А., Шипко В.В., Трифонов П.А. Итерационный алгоритм восстановления цифровых изображений с адаптивным обнаружением импульсных помех // Матер. XIX Междунар. науч. конф. «Радиолокация навигация связь». Воронеж: Изд-во НПФ «САКВОЕЕ» ООО, 2013. Т. 1. С. 182–189.

11. Самойлин Е.А. Алгоритмы оценивания импульсного шума в задачах цифровой фильтрации оптических изображений // Оптический журнал. 2006. Т. 73. № 12. С. 42–46.

12. Калинин П.В., Сирота А.А. Моделирование аппликативных искажений с различной степенью прозрачности и случайной формой // Цифровая обработка сигналов. 2013. №1. С. 28-33.

## METHOD OF THE DISTINGUISHING CASUAL SIGNAL BY MUCH COMPONENTS IMAGES AND IMPULSE NOISE BY EXPLOITING INTERCHANNEL REDUNDANCY

*Samoylin E.A., Shipko V.V.*

The offered method of the distinguishing casual signal much components images and impulse noise, founded on calculation and collation interchannel gradient estimation. The presented results of the numerical studies of the method, have shown that reduction of the number mistake distinguishing signal and hindrances in contrast with the known rank algorithm of the processing much components images.

## Уважаемые коллеги!

Для тех, кто не успел оформить подписку на второе полугодие 2014 года через ОАО «Роспечать», сохраняется возможность приобретения журналов непосредственно в редакции по адресу: г. Москва, ул. Авиамоторная, дом 8, Научный Центр МГУСИ, ком. 612. Российское научно-техническое общество радиотехники, электроники и связи им. А.С. Попова, метро «Авиамоторная», или оформить Заказ в соответствии с требованиями, выставленными на сайте журнала: [www.dsra.ru](http://www.dsra.ru).

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

**E-mail:** [rntores@mail.ru](mailto:rntores@mail.ru)

УДК 621.397.42: 621.391.837.1: 621.372.544: 621.3.011.73

**МЕТОД ПОВЫШЕНИЯ РАЗРЕШАЮЩЕЙ СПОСОБНОСТИ ИЗОБРАЖЕНИЙ**

*Дрынкин В.Н., начальник сектора ФГУП ГосНИИ Авиационных систем, г. Москва, e-mail: drynkin@gosniias.ru;*  
*Царева Т.И., к.б.н., старший научный сотрудник ФГУП ГосНИИ Авиационных систем,*  
*e-mail: tsareva@gosniias.ru*

**Ключевые слова:** разрешающая способность, твердотельные матричные фотоприемники, гексагональная дискретизация, трехмерный интерполирующий фильтр, рекурсивно-нерекурсивные фильтры.

**Введение**

Разрешающая способность (РС) характеризует возможности видеосистемы разделять воспроизводимые мелкие детали наблюдаемой сцены и является одним из основных показателей качества изображения. Пространственная РС определяется количеством элементов твердотельных матричных сенсоров (ПЗС, КМОП, фотодиодных матриц InGaAs, микроболометров a-Si и др.), используемых в качестве приемников излучения.

Часто применение видеокамер с большим числом элементов матрицы бывает затруднительно по разным причинам: из-за ограничений применяемых стандартов, из экономических соображений, из-за технологических ограничений и т.п.

Альтернативным подходом повышения разрешения формируемых изображений является использование методов восстановления изображений высокого разрешения (ВР) по последовательности соседних кадров низкого разрешения (НР). Здесь можно выделить два основных направления: методы повышения РС видеокамер на основе применения нескольких матриц с электронной шивкой изображений и методы, основанные на использовании субпиксельного сдвига в соседних кадрах изображения одного и того же объекта.

Среди представителей первого направления практический интерес представляют методы на основе двух матриц ПЗС. Так, в устройстве [1] для получения телевизионных изображений световой поток делится на две части по горизонтали с помощью волоконно-оптического преобразователя и подается на две матрицы ПЗС.

В другой системе [2] для формирования видеосигнала применены две матрицы с зазорами между фоточувствительными элементами, которые взаимно располагаются так, что элементы одной матрицы смещены по горизонтали относительно элементов второй матрицы на половину их ширины. Световой поток от объектива делится между матрицами с помощью полупрозрачного и отражающего зеркал.

В телекамере [3] изображение формируется также с помощью двух матриц ПЗС, а световой поток делится между ними с помощью анаморфотного объектива, изменяющего масштаб изображения в одном направлении,

*Предложен метод повышения разрешающей способности изображений, формируемых твердотельными матричными фотоприемными устройствами. Метод основан на использовании преимуществ гексагональной дискретизации двух соседних кадров с последующей пространственно-временной интерполяцией с помощью трехмерного рекурсивно-нерекурсивного фильтра нижних частот. Приведены результаты полунатурного моделирования с использованием реальных изображений.*

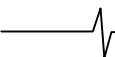
и светоделителя призмного типа.

При всех своих достоинствах и недостатках, отличающих их друг от друга, эти методы характеризуются тем, что повышение РС камеры осуществляется только в одном направлении по горизонтали.

Существуют методы повышения РС в двух направлениях по горизонтали и вертикали, основанные на использовании субпиксельного сдвига в соседних кадрах изображения одного и того же объекта. В литературе такой подход часто называют «супер-разрешением». Главным достижением этого подхода признается то, что он имеет низкую стоимость и позволяет использовать существующие системы НР. Этот метод находит широкое применение в различных областях, в том числе в медицине, в спутниковых изображениях, в видео приложениях и т.п. В работе [4] содержится технический обзор существующих подходов. В [5] рассмотрены теоретические аспекты различных способов получения супер-разрешения.

Однако при использовании метода супер-разрешения возникает целый ряд неоднозначно решаемых вопросов: как получать смещенные изображения, сколько кадров необходимо, какова будет кратность увеличения разрешения, как определить смещение между кадрами, как бороться с шумом, возникающим от дублирования информации, не теряя получаемую детальность и т.д.

Другим подходом к увеличению РС вдоль обеих осей декартовых координат является способ с использованием светоделительной призмы и двух матриц, устанавливаемых со сдвигом относительно друг друга по горизонтали – на величину, равную половине расстояния между элементами, и по вертикали – на величину, равную половине межстрочного промежутка, в результате чего в блоке памяти образуется информационное поле с шахматной структурой, на основе которого с помощью обработки в виде гребенчатой фильтрации и полной апертурной коррекции формируется телевизионное изображение высокой четкости [6]. Существенным недостатком данного способа является применение для формирования окончательного изображения простейшей одномер-



ной интерполяции, заключающейся в усреднении видеосигнала  $n$ -ой строки кадра первого изображения с усредненным видеосигналом  $(n-1)$ -ой и  $n$ -ой строк кадра второго изображения, которая, как известно, не обеспечивает высокую степень подавления помех дискретизации и высокочастотных шумов [7, 8].

В данной работе предлагается метод повышения РС изображений по вертикали и горизонтали, свободный от перечисленных выше недостатков. Он основан на способе, использующем диагональный субпиксельный сдвиг матричного изображения в соседних кадрах и обеспечивающем повышение РС видеосистем за счет трехмерной интерполяции этих кадров [9].

### Описание способа повышения разрешающей способности

Согласно этому способу четные и нечетные кадры НР снимают со сдвигом по диагонали на полпикселя друг относительно друга, после чего размеры этих кадров увеличивают путем введения в их матричную структуру промежуточных нулевых столбцов и строк. В результате образуется пространственно-временная решетка двух кадров с шахматной структурой расположения отсчетов (рис. 1) [10, 11]. Затем производят формирование последовательности выходных сигналов путем одновременного или последовательного считывания данных каждого увеличенного нечетного и соответствующего увеличенного четного кадров, которые обрабатывают с помощью трехмерного интерполирующего пространственно-временного фильтра нижних частот (ФНЧ), трехмерная частотная характеристика которого имеет область пропускания в виде октаэдра. На выходе фильтра формируется кадр изображения удвоенного пиксельного размера, в котором за счет трехмерной интерполяции полностью восстановлены все нулевые элементы [10, 11]. При этом обеспечивается практически двукратное повышение РС по двум направлениям.

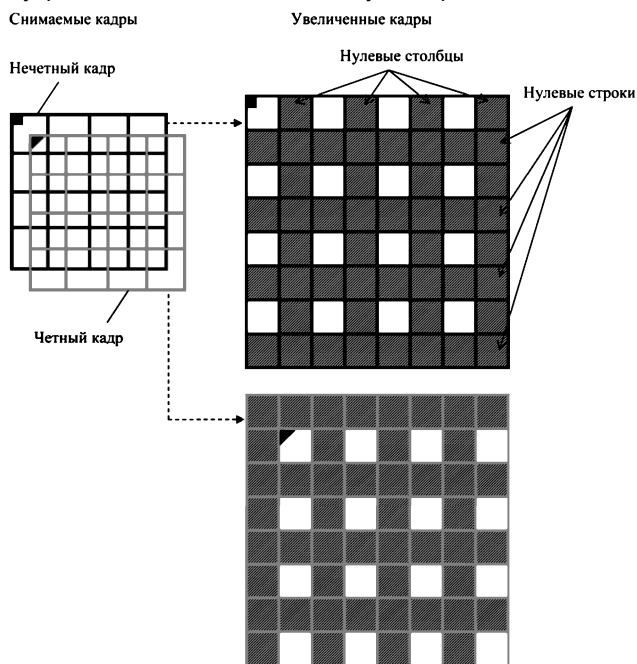


Рис. 1. Формирование решетки двух кадров с шахматной структурой расположения отсчетов

### Формирование кадров с субпиксельным сдвигом

Различные области применения изображений повышенного разрешения определяют способы реализации двух смещенных по диагонали на полпикселя кадров. Рассмотрим некоторые варианты.

#### С помощью одной матрицы

Формирование двух кадров, смещенных на полпикселя по диагонали, можно осуществлять с помощью одной матрицы за счет естественного вибрирования носителя (камеры) аналогично тому, как это делается в вышеупомянутых методах получения супер-разрешения. Существенным недостатком данного способа является необходимость накопления нескольких кадров, сложности определения смещений между кадрами, необходимость дополнительной интерполяции для ослабления шумов, возникающих из-за неточного смещения кадров на половину пикселя и другие проблемы, характерные для методов супер-разрешения.

Более стабильных результатов получения субпиксельного сдвига двух кадров можно достичь, используя управляемый механический сдвиг, применяемый, например, в так называемых джиттер-камерах [12]. На рис. 2 показаны два возможных варианта. Сдвиг соседних кадров осуществляется с помощью управляемых микро-актюаторов, которые могут воздействовать как на саму матрицу (рис. 2, а), так и на вспомогательное оптическое устройство типа наклонной стеклянной пластины, устанавливаемой между объективом и неподвижной матрицей (рис. 2, б).

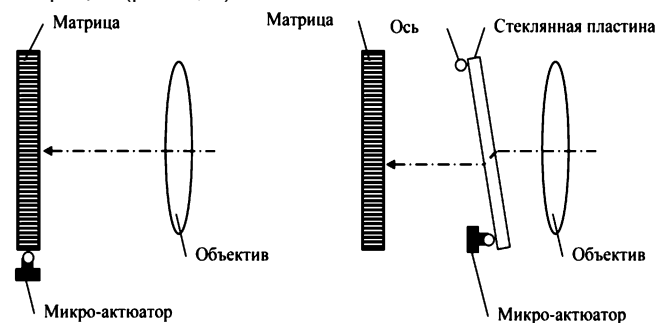


Рис. 2. Примеры механической реализации диагонального субпиксельного сдвига кадров

Точность субпиксельного сдвига в таких камерах может достигать 0,1 пикселя [12].

Более точных результатов можно достичь, используя управляемый электронный субпиксельный сдвиг, например, в камерах, использующих явление биннинга. Известно, что биннинг (т.е. суммирование зарядов смежных ячеек матрицы) применяется с целью увеличения чувствительности камеры, повышения отношения сигнал-шум и возможности увеличения частоты кадров. Однако этот процесс сопровождается уменьшением разрешения изображения. С помощью предлагаемого метода с одной стороны можно повысить разрешение камеры, использующей биннинг, с другой стороны применение биннинга позволяет упростить проблему формирования двух соседних диагонально смещенных друг относительно друга кадров с помощью одной матрицы. На рис. 3 показан пример организации считывания информации из матрицы для двух соседних кадров.

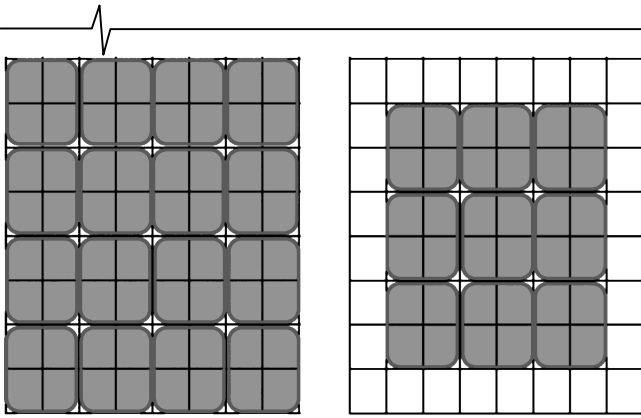


Рис. 3. Организация процедуры считывания информации с матрицы с помощью биннинга для двух соседних кадров

Особенностью рассмотренных методов формирования диагонального субпиксельного сдвига в соседних кадрах с помощью одной матрицы является то, что для последующей обработки информации необходима реализация трехмерного интерполирующего ФНЧ с осуществлением задержки на кадр.

**С помощью двух матриц**

Применение двух одинаковых матриц позволяет предложить схему формирования двух кадров с диагональным субпиксельным сдвигом, изображенную на рис. 4. Деление светового потока между матрицами осуществляется с помощью полупрозрачного и непрозрачного зеркал (рис. 4, а). Матрицы располагаются друг относительно друга так, чтобы их фоточувствительные элементы формировали кадры сцены, смещенные на полпикселя по диагонали (рис. 4, б). При такой схеме в каждый момент времени мы имеем два диагонально смещенных кадра, что позволяет обойтись без устройства задержки на кадр и, тем самым, упростить реализацию трехмерного интерполирующего фильтра.

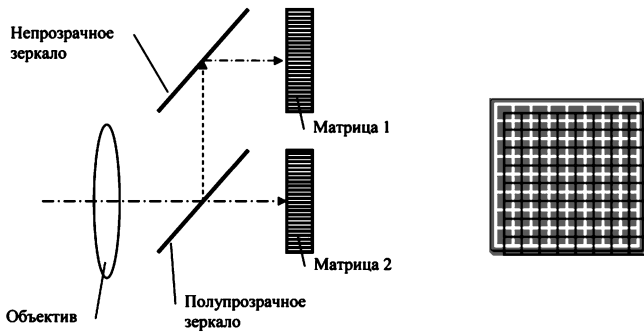


Рис. 4. Схема размещения двух матриц, в одной камере для формирования диагонального субпиксельного сдвига кадров

Для одновременного получения двух смещенных по диагонали кадров можно использовать две матрицы, чувствительные к разным длинам волн спектра электромагнитного излучения и, таким образом, использовать предлагаемый метод для повышения разрешающей способности двухзональной системы, например, телевизионно-тепловизионной [8].

**Результаты синтеза трехмерного ФНЧ**

Согласно теории дискретизации восстановление изображения по его отсчетам сводится к синтезу многомерного интерполирующего фильтра, область пропускания частотной характеристики которого согласована с областью существования спектра изображения [11]. В случае

гексагональной дискретизации с шахматным расположением отсчетов, учитывая анизотропию спектров реальных изображений [11, 13, 14], определим область пропускания частотной характеристики трехмерного интерполирующего ФНЧ в виде октаэдра:

$$D_0 : |v_x| + |v_y| + |v_t| = a, \tag{1}$$

где  $a \leq 1$ ;  $v_x, v_y, v_t$  – соответствующие нормированные продольные, поперечные и временные частоты.

Для синтеза фильтра с областью пропускания (1) воспользуемся методом многомерной рекурсивно-нерекурсивной фильтрации [15, 16], который был специально разработан для синтеза фильтров с произвольной областью пропускания частотной характеристики. Учитывая, что кадр ВР формируется из двух соседних кадров НР ограничимся трёхмерным фильтром первого порядка с фиксированным нулём передаточной функции  $z_0 = -1$  на основе Чебышевского одномерного аналогового прототипа, имеющего один вещественный полюс  $w_p$ .

Синтезируемый фильтр представляет собой каскадное включение трехмерного  $K[v_t, \phi_t(v_x, v_y)]$ , двумерного  $K[v_y, \phi_y(v_x)]$  и одномерного  $K(v_x)$  звеньев:

$$K(v_x, v_y, v_t) = K[v_t, \phi_t(v_x, v_y)] K[v_y, \phi_y(v_x)] K(v_x). \tag{2}$$

Конфигурация области пропускания синтезируемого фильтра в направлении временных частот  $v_t$  полностью определяется трехмерным рекурсивно-нерекурсивным звеном

$$K[v_t, \phi_t(v_x, v_y)] = \frac{0,5(1 + e^{-j\pi v_t})}{1 - \beta(v_x, v_y) e^{-j\pi v_t}} [1 - \beta(v_x, v_y)], \tag{3}$$

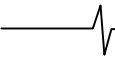
$$\text{где } \beta(v_x, v_y) = \frac{\text{ctg} \frac{\pi(a - v_x - v_y)}{2} + w_p}{\text{ctg} \frac{\pi(a - v_x - v_y)}{2} - w_p} \tag{4}$$

представляет собой пространственно-частотную характеристику (ПЧХ) двумерной нерекурсивной цепи обратной связи. Для получения практически реализуемой структуры примем  $a = 0,8$  и, учитывая, что Чебышевский одномерный аналоговый прототип при неравномерности частотной характеристики в полосе пропускания  $\delta = 1$  дБ имеет  $w_p = -1,9652267$  [17], аппроксимируем (4) двумерным тригонометрическим рядом:

$$\hat{\beta}(v_x, v_y) = \gamma 0,656 - 2 \cdot 0,156(\cos \pi v_x + \cos \pi v_y) - 4 \cdot 0,109 \cos \pi v_x \cos \pi v_y. \tag{5}$$

Коэффициенты этого ряда определяются с помощью двумерного обратного быстрого преобразования Фурье (ОБПФ) относительно (4). Коэффициент  $\gamma = 0,81$  выбирается из соображений обеспечения устойчивости синтезируемого фильтра.

Конфигурация области пропускания синтезируемого



фильтра в плоскости изображения  $(v_x, v_y)$  определяется двумерным рекурсивно-нерекурсивным звеном:

$$K[v_y, \phi_y(v_x)] = \frac{0,5(1 + e^{-j\pi v_y})}{1 - \beta(v_x)e^{-j\pi v_y}} [1 - \beta(v_x)], \quad (6)$$

где  $\beta(v_x, v_y) = \frac{\text{ctg} \frac{\pi(a - v_x)}{2} + w_p}{\text{ctg} \frac{\pi(a - v_x)}{2} - w_p}$ . (7)

Для получения практически реализуемой структуры этой цепи аппроксимируем (7) одномерным тригонометрическим рядом с учетом  $a = 0.8$  и  $w_p = -1,9652267$ :

$$\hat{\beta}(v_x) = 0,534 + 2 \cdot (-0,277) \cos \pi v_x + 2 \cdot (-0,133) \cos 2\pi v_x + 2 \cdot (0,0025) \cos 3\pi v_x + 2 \cdot (0,0151) \cos 4\pi v_x. \quad (8)$$

Одномерное звено  $K(v_x, v_y)$  выражения (2) задает частоту среза синтезируемого фильтра в направлении частот  $v_x$  и для фильтра первого порядка имеет вид

$$K(v_x) = \frac{0,5(1 + e^{-j\pi v_x})}{1 - \beta e^{-j\pi v_x}} [1 - \beta], \quad (9)$$

где с учётом сделанных выше допущений

$$\beta = \frac{\text{ctg} \frac{\pi a}{2} + w_p}{\text{ctg} \frac{\pi a}{2} - w_p} = -0,716. \quad (10)$$

Результирующая ПЧХ трёхмерного рекурсивно-нерекурсивного ФНЧ показана на рис. 5 в виде поверхности уровня  $K(v_x, v_y, v_t) = 0,8$  для положительного октанта области нормированных частот. Степень практического соответствия требуемой и получаемой конфигураций области пропускания будет определяться точностью аппроксимации частотных характеристик рекурсивных фильтров обратных связей и, следовательно, протяженностью импульсных характеристик этих фильтров [15].

Нетрудно видеть (рис. 5), что даже при простейших аппроксимациях цепей обратных и прямых связей тригонометрическими рядами (5) и (8) область пропускания синтезированного фильтра довольно точно соответствует требуемой (1).

Обобщенная структурная схема синтезированного трехмерного интерполирующего ФНЧ с учетом (3), (5), (6), (8-10) при формировании диагонально смещенных

кадров с помощью одной матрицы показана на рис. 6 а, где  $e^{-j\pi v_t}$ ,  $e^{-j\pi v_y}$  и  $e^{-j\pi v_x}$  представляют собой устройства задержки на кадр, строку и элемент изображения, соответственно.

Здесь:  $S_1(v_x, v_y, v_t)$  и  $S_2(v_x, v_y, v_t)$  – Фурье-образы первого и второго диагонально смещенных кадров НР,  $S_{BP}(v_x, v_y, v_t)$  – Фурье-образ кадра ВР. При формировании диагонально смещенных кадров с помощью двух матриц структурная схема ФНЧ показана на рис. 6 б. В этом случае, как указывалось выше, память на кадр  $e^{-j\pi v_t}$  отсутствует.

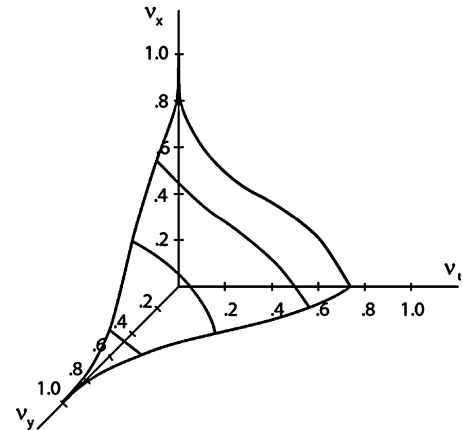


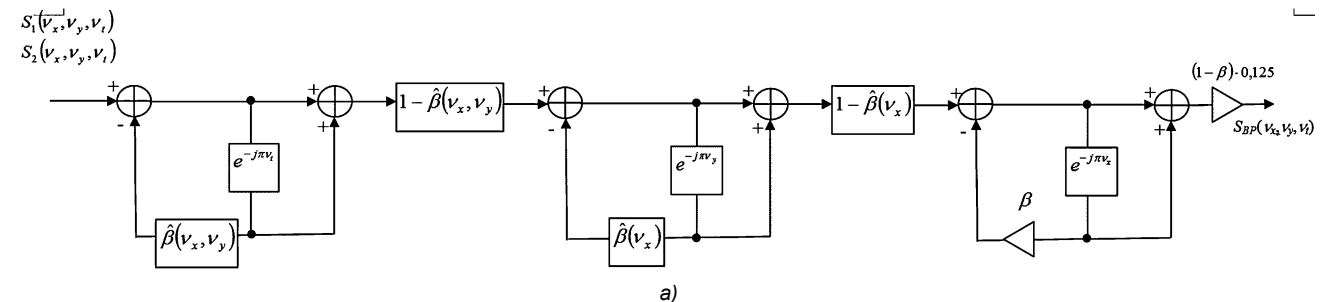
Рис. 5. ПЧХ синтезированного трехмерного ФНЧ

Синтезированный фильтр имеет ряд достоинств.

Во-первых, важным достоинством синтезированного трехмерного ФНЧ является его контролируемая устойчивость, которая обеспечивается соответствующим проектированием цепей прямых и обратных связей  $\hat{\beta}(v_x, v_y)$  (5),  $\hat{\beta}(v_x)$  (8). В данной работе устойчивость синтезированного ФНЧ детально не исследуется, поскольку этому вопросу посвящена специальная статья [18].

Во-вторых, область пропускания ПЧХ синтезированного фильтра согласована со спектрами реальных изображений, что обеспечивает лучшее по сравнению с одномерным аналогом [6] качество интерполяции изображения повышенного разрешения. Сравнение трехмерной интерполяции с одномерными вариантами проведено в [8, 14, 20].

В-третьих, рекурсивно-нерекурсивная структура фильтра имеет сравнительно небольшие аппаратные затраты (при реализации требуется память всего на один кадр), что обеспечивает его эффективную реализацию в реальном времени [19, 20].





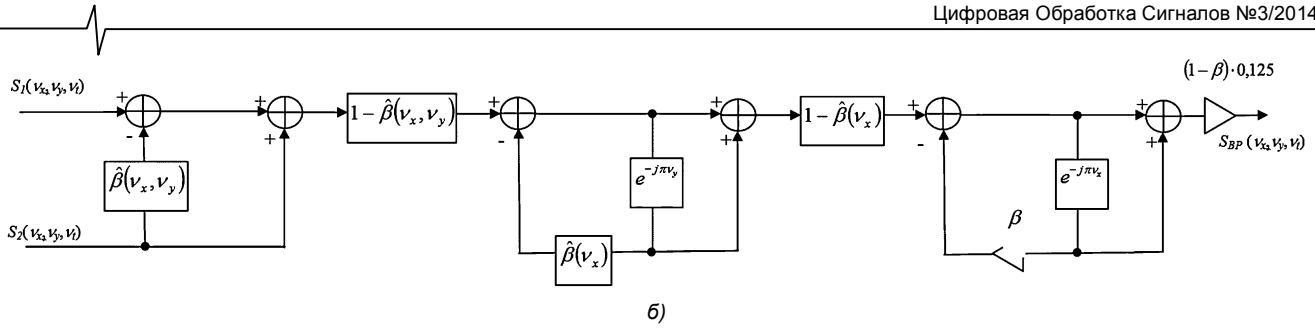


Рис. 6. Обобщенная структурная схема синтезированного трехмерного ФНЧ

**Тестирование метода**

Методами математического полунатурного моделирования с использованием реальных изображений была продемонстрирована возможность повышения разрешающей способности изображений с помощью предложенного метода.

На рис. 7 а, 8 а, 9 а приведены примеры синтезированных предложенным методом изображений повышенного разрешения. Для сравнения на рис. 7 б, 8 б, 9 б показаны по одному из двух исходных кадров НР, смещенных друг относительно друга на полпикселя по диагонали.

Из рис. 7 видно, что в результате повышения разрешения стали видны фрагменты железнодорожного полотна, а на рис. 8 на фоне общего повышения четкости мелких деталей автомобиля уверенно читается государственный регистрационный номер.



Рис. 7. Пример 1 повышения разрешения

Проведенные исследования показали, что предлагаемый метод позволяет получить близкое к двукратно-

му повышение разрешения. Для получения количественной оценки повышения разрешающей способности изображения была проведена серия экспериментов со стандартными измерительными штриховыми мирами.



Рис. 8. Пример 2 повышения разрешения

Разрешение определялось методом визуальных экспертных оценок на основе критерия, согласно которому разрешение определяется по группе штрихов с самой узкой шириной полос, разрешаемых по меньшей мере 80% всех наблюдателей. При этом согласно критерию оценки разрешающей способности по частотно-контрастным характеристикам для анализа использовались изображения групп штрихов с относительным контрастом, определяемым разницей между максимальной и минимальной интенсивностью, отнесенной к максимальной интенсивности, не менее 0,2. Группа штрихов считается разрешенной, если по всей длине чередующихся темных и светлых полос при сохранении их одинаковой ширины визуально ощущается разница в уровнях серого тона.

Анализ полученных результатов показал, что в среднем разрешение повышается в 1,7 раза.

На рис. 9 приведен пример, демонстрирующий эту величину. На этом рисунке показаны фрагменты изображений стандартной штриховой миры Фуко. Каждая группа штрихов миры состоит из чередующихся черных и белых полос одинаковой ширины, ориентированных в четырех разных направлениях. Ширина штрихов убывает с увеличением номера группы по закону геометрической прогрессии со знаменателем, равным 0,94. На фрагменте кадра НР (рис. 9 б) видно, что по всей длине штрихов с сохранением одинаковой ширины темных и светлых полос уверенно различается группа, состоящая из 6 штрихов (обведено кружком) с шириной полосы 1,0 мм, в то время как в изображении повышенного разрешения (рис. 9 а) различается группа из 10 штрихов с шириной полосы 0,6 мм.

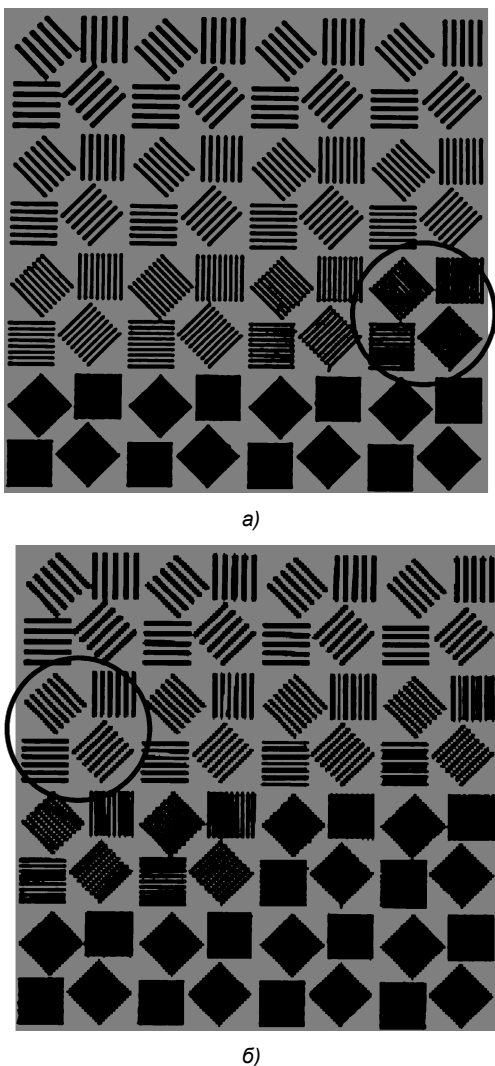


Рис. 9. Пример 3 повышения разрешения

## Заключение

Разработан метод, позволяющий повысить РС по двум направлениям плоскости изображения за счет двукратного увеличения пиксельного размера кадра с помощью трехмерной интерполяции двух соседних кадров НР, смещенных относительно друг друга на полпикселя по диагонали.

Метод опробован на целом ряде тестовых и реальных изображений. Полученные результаты являются

предварительными и требуют дальнейших исследований, тем не менее уже сейчас они позволяют рассматривать предложенный метод в качестве основы для повышения РС телевизионных, тепловизионных и др. видеосистем, в том числе многозональных.

## Литература

1. Гуйо Л., Рикодо Ж., Ружо А. Устройство для получения телевизионных изображений с помощью матриц с зарядовой связью и передающая система с подобным устройством. Заявка Франции № 2476949, H04N 5/32; H05G 1/60, 1/64. Опубл. 28.08.81.
2. Hoagland K.A. Charge-coupled device video-signal-generating system. US Patent № 4038690, H04N 3/14, 358/213; 357/24; 357/30. Filed: Jan. 21, 1976. Publ. July 26, 1977.
3. Смелков В.М. Метод повышения разрешающей способности телекамеры для криминалистической диагностики // журнал "Специальная техника" Дата размещения: 12.04.2006 [Электронный ресурс]. URL: <http://www.bnti.ru/showart.asp?aid=625&lv=01.03>. (дата обращения: 03.04.2014).
4. Park S.Ch., Park M.K., Kang M.Gi. Super-resolution image reconstruction: a technical overview // Signal Processing Magazine. IEEE. 2003. May. Volume 20. Issue 3. Pp. 21–36.
5. Katsaggelos A., Molina R., Mateos J. Super Resolution of Images and Video. Synthesis Lectures on Image, Video and Multimedia Processing. / Editor A.C. Bovik. Morgan&Claypool Publishers. 2007. 134 pp.
6. Виленчик Л.С., Курков И.Н., Разин А.И., Розвал Я.Б. Способ получения телевизионных изображений высокой четкости в камере на обычных ПЗС и устройство для реализации этого способа. Патент РФ № 2143789, H04N 5/335, 5/225. Заявл. 23.01.1998. Опубл. 27.12.1999.
7. Бородянский А.А., Дрынкин В.Н. Вертикально-временная фильтрация в системах телевидения повышенной четкости / Рязань. 1986. 15 с. – Деп. в ЦНТИ «Имформсвязь» 24.03.87, № 1068-св.
8. Дрынкин В.Н., Фальков Э.Я., Царева Т.И. Эффективность двухзональной бортовой авиационно-космической системы формирования комбинированного изображения // Техническое зрение в системах управления – 2013. / Труды научно-технической конференции. – М.: ИКИ РАН, 12-14 марта 2013 г. (в печати).
9. Дрынкин В.Н., Царева Т.И. Способ повышения разрешающей способности видеосистем. Заявка на изобретение № 2014103333(005183) от 03.02.2014. H04N 3/14, 5/335. / Заявитель – ФГУП ГосНИИАС.
10. Бородянский А.А. Оптимальная дискретизация движущихся изображений // Электросвязь. 1983. № 3. С. 35-39.
11. Бородянский А.А. Гипертреугольная дискретизация n-мерных сообщений // Радиотехника. 1985. № 4. С. 49-52.
12. Ben-Ezra M., Zomet A., Nayar S.K. Video Super-Resolution Using Controlled Subpixel Detector Shifts // IEEE Transactions on Pattern Analysis and Machine Intelligence. June 2005. Vol. 27. No. 6. P. 977-987.

13. Цифровое кодирование телевизионных изображений; под ред. И.И. Цуккермана. М.: Радио и связь, 1981. 240 с.

14. Дрынкин В.Н., Фальков Э.Я., Царева Т.И. Формирование комбинированного изображения в двухзональной бортовой авиационно-космической системе // Техническое зрение в системах управления – 2012. Труды научно-технической конференции. / М.: ИКИ РАН, 14-16 марта 2012 г. – С. 33-39.

15. Бородянский А.А., Дрынкин В.Н. Синтез многомерных рекурсивно-нерекурсивных фильтров // Радиотехника. 1986. № 4. С. 47-51.

16. Drynkin V.N. Real-Time design of N-dimensional digital filters for image processing // Digital Photogrammetry and Remote Sensing'95; editor E.A. Fedosov. St. Petersburg, 1995. P. 240-249.

17. Альбац М.Е. Справочник по расчету фильтров и линий задержки. М.–Л.: Госэнергоиздат, 1963. 200 с.

18. Бородянский А.А., Дрынкин В.Н. Устойчивость многомерных рекурсивно-нерекурсивных фильтров // Радиотехника. 1988. № 3. С. 37-38.

19. Бондаренко А.В., Докучаев И.В., Дрынкин В.Н.,

Царева Т.И., Бондаренко М.А. Аппаратная реализация трёхмерного фильтра // Техническое зрение в системах управления – 2013. Труды научно-технической конференции. / М.: ИКИ РАН, 12-14 марта 2013 г. (в печати).

20. Бондаренко А.В., Бондаренко М.А., Дрынкин В.Н., Докучаев И.В., Ядчук К.А. Пространственно-временная фильтрация движущихся изображений // Техническое зрение в системах управления – 2014. Труды научно-технической конференции. / М.: ИКИ РАН, 18-20 марта 2014 г. (в печати)

## RESOLUTION INCREASING METHOD OF IMAGES

*Drynkin V.N., Tsareva T.I.*

Method for resolution increasing of images formed by solid-state photodetector arrays is proposed. The method is based on the use of advantages of the hexagonal sampling of two adjacent frames with subsequent spatiotemporal interpolation using a three-dimensional-recursive-nonrecursive lowpass filter. The results of half-scale simulation with the use of the real images are given.

### Уважаемые коллеги!

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

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

Научно-технический журнал «Цифровая обработка сигналов» включен в Список изданий, рекомендуемый ВАК РФ для публикации результатов научных исследований соискателями ученой степени доктора и кандидата технических наук в области радиотехники, связи, вычислительной техники, электроники, информационно-измерительных и управляющих систем.

### Планируемые сроки издания отдельных номеров журнала:

- № 4 – декабрь 2014 г. Тематический выпуск: «ЦОС в радиотехнике и системах телекоммуникаций».
- № 1 – март 2015 г. Тематический выпуск: «ЦОС в информационно-измерительных системах»
- № 2 – июнь 2015 г. Тематический выпуск: «Теория и методы цифровой обработки сигналов» (по материалам международной научно-технической конференции «Цифровая обработка сигналов и ее применение – DSPA' 2015».
- № 3 – сентябрь 2015 г. Тематический выпуск: «Цифровая обработка изображений».

### Ориентировочная стоимость рекламных услуг:

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

С наилучшими пожеланиями, зам. главного редактора  
д.т.н., профессор Витязев Владимир Викторович, телефон 8-903-834-81-81.

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

## ПРЕДВАРИТЕЛЬНОЕ СОВМЕЩЕНИЕ ИЗОБРАЖЕНИЙ И МЕТОДЫ ОЦЕНКИ КАЧЕСТВА СОВМЕЩЕНИЯ

*Новиков А.И., доцент кафедры «Высшая математика» ФГБОУ ВПО «Рязанский государственный радиотехнический университет», e-mail: novikovanatoly@yandex.ru;*

*Ефимов А.И., магистрант кафедры «Электронные вычислительные машины» ФГБОУ ВПО «Рязанский государственный радиотехнический университет», e-mail: lexie62rus@mail.ru*

**Ключевые слова:** реальное и виртуальное изображения, контуры объектов, ключевые точки, качество совмещения.

### Введение

Необходимость решения задачи совмещения изображений и использования полученных результатов возникает во многих областях науки и техники. Это картография, дистанционное зондирование Земли, многоспектральные системы технического зрения (СТЗ) летательных аппаратов (ЛА), зрение робототехнических систем, системы автоматического роботизированного управления. Настоящее исследование ориентировано на применение в многоспектральных СТЗ ЛА. В современной авиации техническое зрение решает такие актуальные задачи, как корреляционно-экстремальная навигация, обнаружение и сопровождение целей, визуализация изображения кабинного пространства для пилота и др.

Разработка и создание систем комбинированного видения (CVS) и улучшенного видения (ESVS) в настоящее время является исключительно важной задачей. Это обусловливается требованиями повышения безопасности полетов в сложных метеоусловиях, при плохой видимости, в ночное время. Интенсивное развитие и внедрение систем, обеспечивающих совмещение реальных изображений, получаемых от датчиков различной физической природы, и изображений, синтезированных по цифровой карте местности, позволяет создавать и отображать комбинированное изображение и, как следствие, обеспечивать управление ЛА в сложных условиях.

В качестве источников информации в таких системах могут использоваться: видео датчики, инфракрасные (ИК) датчики различных диапазонов, миллиметровые радары (МР), лазерные локаторы (ЛЛ), базы данных рельефа местности вдоль маршрутов полёта, базы данных аэропортов и объектов взлётно-посадочной полосы (ВПП), навигационные параметры и др.

Сформированная системой ESVS оперативная графическая информация далее представляется пилоту в реальном времени на соответствующем устройстве отображения – индикаторе на лобовом стекле (ИЛС) или многофункциональном индикаторе (МФИ), представляющем собой компьютерный дисплей, входящий в состав бортовой системы отображения информации.

Совмещение реального и синтезированного изобра-

*Приведен алгоритм предварительного совмещения реального изображения и изображения, синтезированного по цифровой карте местности, предназначенный для реализации в бортовых вычислительных комплексах летательных аппаратов. Применение алгоритма позволяет локализовать в пространстве шести параметров точку с истинными координатами воздушного судна. Одновременно предложены методы оценки качества совмещения реального и виртуального изображений.*

жений в бортовых СТЗ является одной из сложнейших задач по нескольким причинам. Одна из основных заключается в ошибках определения текущих координат ЛА как материальной точки в воздушном пространстве (широты  $\lambda$ , долготы  $\varphi$  и высоты  $h$ ), а также в ошибках определения ориентации ЛА как протяженного объекта в пространстве. К ним относятся ошибки в измерении параметров курса  $\psi$ , тангажа  $\theta$  и крена  $\gamma$ .

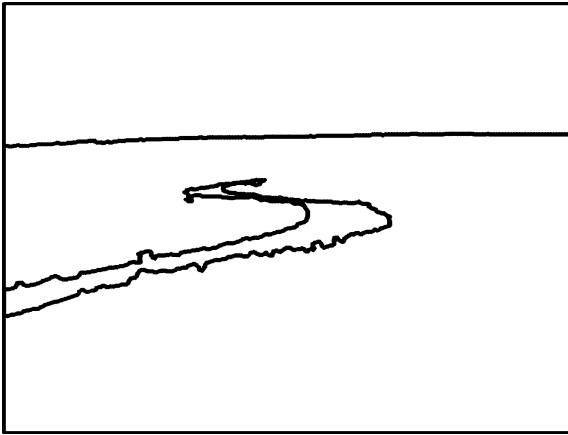
Для совмещения реального и синтезированного изображений приходится генерировать некоторое множество ракурсов виртуальной карты местности. Под ракурсом будем понимать конкретный набор значений шестимерного арифметического вектора  $\mathbf{v} = (x, y, h, \psi, \theta, \gamma)$  – параметров позиционирования ЛА. Параметры позиционирования ЛА являются исходными данными для формирования синтезированного изображения по цифровой карте местности.

К ошибкам позиционирования могут добавляться еще и геометрические искажения контуров объектов, привносимые в обработанные изображения на этапах выделения границ перепада яркостей и формирования замкнутых контуров. Разная природа реального и синтезированного изображений является еще одной причиной, затрудняющей решение задачи совмещения изображений. Гладким контурам береговой линии водоема на синтезированном изображении могут отвечать линии с большей степенью изменчивости (колебания) на обработанном реальном изображении.

Пример пары таких изображений – реального (рис. 1а) и синтезированного по цифровой карте местности (рис. 1б), содержащего контуры линии горизонта и реки, приведен на рис. 1.

Синтезированное изображение получено из цифровой карты местности по навигационным параметрам, которые содержат неизвестные нам ошибки. Поэтому простое наложение изображений, представленных на рис. 1, обнаруживает существенное несовпадение кон-

туров основных объектов (линии горизонта и контуров реки). Результат наложения изображений приведен на рис. 2 (черным цветом показано реальное изображение, серым – синтезированное). Сравнение линий горизонта реального и виртуального изображений позволяет утверждать, что они различаются по углам крена; контуры реки на изображениях смещены в плоскости Оху, а это означает, что оставшиеся пять навигационных параметров также могут содержать ошибки.



а)



б)

Рис. 1. Обработанное реальное (а) и отвечающее ему синтезированное (б) изображения

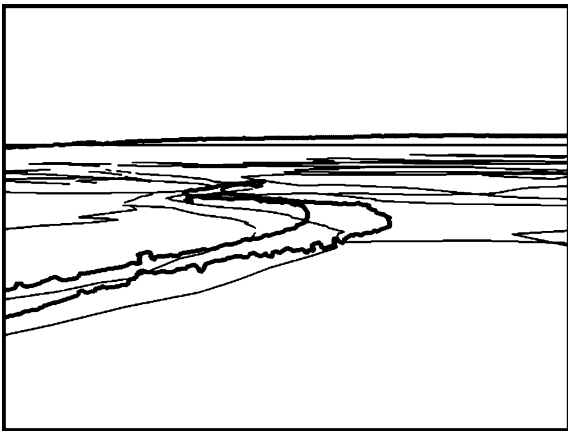


Рис. 2. Результат наложения реального и синтезированного изображений, представленных на рис. 1

### Алгоритмы совмещения изображений

Наиболее известными и часто применяемыми в практике совмещения изображений являются корреля-

ционные алгоритмы совмещения [1]. Они дают достаточно хорошие результаты совмещения [2]. Однако применение этих методов в бортовых СТЗ летательных аппаратов нереально из-за их высокой вычислительной сложности, исключающей возможность решения задач в реальном времени. Естественным в этих условиях является переход от сравнения самих изображений к сопоставлению контуров основных объектов, присутствующих на сравниваемых изображениях, и выработке на этой основе способов совмещения изображений.

Достаточно оптимистичные результаты дает метод совмещения изображений, базирующийся на поиске некоторого множества соответствующих пар ключевых точек на сравниваемых изображениях и нахождении преобразования  $T$ , приводящего одно изображение к плоскости другого. Идея совмещения двух контуров (двух изображений) с помощью матрицы гомографии  $T$  заключается в умножении этой матрицы

$$T = \begin{pmatrix} t_{11} & t_{12} & t_{13} \\ t_{21} & t_{22} & t_{23} \\ t_{31} & t_{32} & 1 \end{pmatrix} \text{ на вектор } X_i = (x_i, y_i, 1)^T, i = \overline{1, n},$$

первые две компоненты которого являются координатами  $i$ -й точки совмещаемого контура (изображения).

Преобразование  $X'_i = T \cdot X_i, i = \overline{1, n}$  выполняет «совмещение» изображений. Например, синтезированное изображение «накладывается» на реальное. Элементы  $t_{ij}$  матрицы гомографии находятся в результате решения соответствующей системы восьми линейных алгебраических уравнений  $G \cdot t = Y$ . Основная матрица  $G$  порядка 8 и вектор свободных членов  $Y$  этой системы формируются из координат четырех пар ключевых точек. Вектор  $t$  – решение системы – содержит первые 8 компонент искомой матрицы гомографии  $T$  [3].

Метод, основанный на использовании матрицы гомографии, обладает достаточно низкой вычислительной сложностью, и, значит, по этому критерию его можно было бы включать в состав математического обеспечения реальных СТЗ ЛА. Недостатком метода является сильная зависимость качества совмещения изображений от того, насколько удачно были выбраны пары ключевых точек на совмещаемых изображениях [4]. Кроме того, в рамках данного метода не решена пока обратная задача – получение оценки вектора ошибок навигационных параметров  $(\Delta x, \Delta y, \Delta h, \Delta \psi, \Delta \theta, \Delta \gamma)$ . Эта оценка необходима для корректировки навигационных параметров.

### Метод ключевой точки

Предлагается подход, позволяющий осуществить предварительное совмещение контуров на тех же идеях ключевых точек. Основные отличия предлагаемого метода от метода, основанного на использовании матрицы гомографии, заключаются в следующем. Во-первых, для реализации этого метода *необходима лишь одна пара ключевых точек*. Во-вторых, он *позволяет* в процессе совмещения изображений (контуров) *получить предварительные оценки пяти компонент  $\Delta x, \Delta y, \Delta h, \Delta \theta, \Delta \gamma$  вектора ошибок навигационных параметров  $v = (x, y, h, \psi, \theta, \gamma)$* . Ошибка  $\Delta \psi$  по курсу в рамках пред-

лагаемого алгоритма не оценивается. Алгоритм предсказания состоит из нескольких шагов. Рассмотрим основные шаги алгоритма и результаты их применения к паре изображений, представленных на рис. 1.

На первом этапе осуществляется аппроксимация контуров основных объектов на реальном и синтезированном изображениях многоугольниками. В нашем случае это контуры реки. Аппроксимирующие многоугольники должны содержать минимальное число вершин (угловых точек), но при этом достаточно точно повторять контуры объекта. Эти требования противоречивы, поэтому к их конкретному описанию предъявляются особые требования.

Поиску угловых структур на векторных изображениях посвящен ряд работ [5, 6]. В [6] для детектирования углов на изображении в каждой точке обрабатываемого контура производится аппроксимация участков контура отрезками, образующими угловую структуру. Параметры пары линейных функций находятся методом наименьших квадратов. Затем находится квадратичная ошибка аппроксимации контура углом, вычисляется значение критерия  $\chi^2$ , с помощью которого находится вероятность  $p_1$  гипотезы о том, что данный участок контура является уголком. После этого оценивается противоположная гипотеза о том, что данный участок является плавно изменяющейся кривой. Для получения ответа на этот вопрос осуществляется аппроксимация рассматриваемого участка контура полиномом третьей степени. Вновь вычисляются квадратичная ошибка аппроксимации, значение критерия  $\chi^2$  и находится вероятность  $p_2$  противоположной гипотезы. Из сравнения полученных значений  $p_1$  и  $p_2$  вероятности гипотез делается вывод, какая из них правдоподобнее.

Приведенные в работе [6] результаты апробации метода на реальных изображениях говорят о хорошем качестве детектирования углов и, в частности, на зашумленных изображениях. Очевидный недостаток метода кроется в высокой вычислительной сложности, не позволяющей использовать его в реальных СТЗ на борту летательного аппарата.

В [3] для решения этой задачи предложен алгоритм, основанный на вычислении косинусов углов между двумя прямыми, проходящими через текущую точку контура. Алгоритм основан на выборе в качестве вершин аппроксимирующего многоугольника тех точек контура, в которых достигается локальный максимум оценок косинусов углов при этих вершинах. Остальные точки замкнутого контура отбрасываются. Оценки косинусов углов вычисляются по нескольким соседним точкам с обеих сторон, причем количество таких точек  $m$  – настраиваемый параметр алгоритма.

В каждой точке контура вычисляется косинус

$$\cos \alpha = \frac{p_1 p_2 + \hat{k}'_1 \hat{k}'_2}{\sqrt{1 + \hat{k}_1^2} \sqrt{1 + \hat{k}_2^2}} \quad (1)$$

угла между прямыми  $L_1 : y = \hat{k}_1 x$  и  $L_2 : y = \hat{k}_2 x$ , проходящими через данную точку и через  $m$  пикселей слева и справа от этой точки. В формуле (1):  $p_1 = 1, p_2 = 1$ ,

$\hat{k}'_1 = \hat{k}_1, \hat{k}'_2 = \hat{k}_2$ , если  $\sum_{i=1}^m x_i > 0$ . Если же  $\sum_{i=1}^m x_i < 0$ , то знаки в равенствах меняются на противоположные, т.е.  $p_1 = -1, p_2 = -1; \hat{k}'_1 = -\hat{k}_1, \hat{k}'_2 = -\hat{k}_2$ .

Оптимальная оценка  $\hat{k}$  углового коэффициента каждой прямой находится как оценка метода наименьших квадратов по формуле

$$\hat{k} = \frac{\sum_{i=1}^m y_i x_i}{\sum_{i=1}^m x_i^2}.$$

В этой формуле  $(x_i, y_i)$  – координаты точек  $A_i, i = \overline{1, m}$  в подвижной системе координат  $Ox$  для углового коэффициента  $k_1$  и  $(x_i, y_i)$  – координаты точек  $B_i, i = \overline{1, m}$  – для углового коэффициента  $k_2$ .

Следующим шагом алгоритма является установление взаимно однозначного соответствия между угловыми точками многоугольников, полученных на предыдущем шаге (поиск пар ключевых точек). Для этого в каждой точке  $M_i$  первого многоугольника вычисляются значение критерия  $\delta_{ij} = \delta(M_i, M_j)$  «похожести» углов

$\alpha_i = \angle M_{i-1} M_i M_{i+1}$  и  $\beta_j = \angle M'_{j-1} M'_j M'_{j+1}$ :

$$\delta_{ij} = \frac{1}{2} \left| \cos \alpha_i - \cos \beta_j \right| + \frac{\left| |\bar{a}_1| - k \cdot |\bar{b}_1| \right|}{|\bar{a}_1| + |\bar{b}_1|} + \frac{\left| |\bar{a}_2| - k \cdot |\bar{b}_2| \right|}{|\bar{a}_2| + |\bar{b}_2|} \quad (2)$$

Здесь  $\bar{a}_1 = \overline{M_i M_{i-1}}, \bar{a}_2 = \overline{M_i M_{i+1}};$   
 $\bar{b}_1 = \overline{M'_j M'_{j-1}}, \bar{b}_2 = \overline{M'_j M'_{j+1}}$   $k = \frac{p_1}{p_2}$ , где  $p_1,$

$p_2$  – периметры первого и второго контуров соответственно. При фиксированном значении индекса  $i$  пройдем все точки  $M'_j$  второго контура из списка  $\{M'_j\}_{j=1}^n$ , вычисляем в каждой из них значение критерия (2) и находим индекс  $j_0$  точки  $M_{j_0}$ , для которой  $\delta(M_i, M_{j_0}) = \min_{j=1, \overline{1, n}} \delta_{ij}$ , т.е.  $j_0 = \arg \min_{j=1, \overline{1, n}} \delta_{ij}$ . Найденная точка

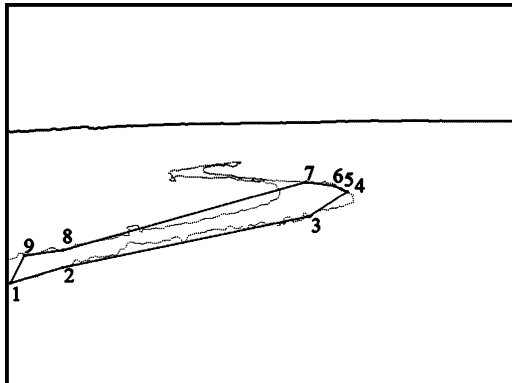
$M_{j_0}$  на втором контуре ставится в соответствие точке  $M_i$  на первом контуре. Это соответствие можно будет считать взаимно однозначным после прохождения всех точек первого контура при условии, что на каком-то шаге алгоритма точка  $M_{j_0}$  не будет поставлена в соответствие еще одной точке первого контура.

Нужно заметить, что на втором контуре может быть найдена более чем одна точка  $M'_j$  с близкими значениями критерия  $\delta_{ij}$ . При этом одна из набора таких точек может быть действительно «идентичной» точке  $M_{i+1}$ , а другие – нет. Для разрешения возникающего противоречия вводится дополнительная проверка «идентичности» сравниваемых точек. Формируются векторы  $\bar{a}'_1 = \bar{a}_1 + \bar{a}_2$  и  $\bar{a}'_1 = \bar{b}_1 + \bar{b}_2$  и вычисляется косинус

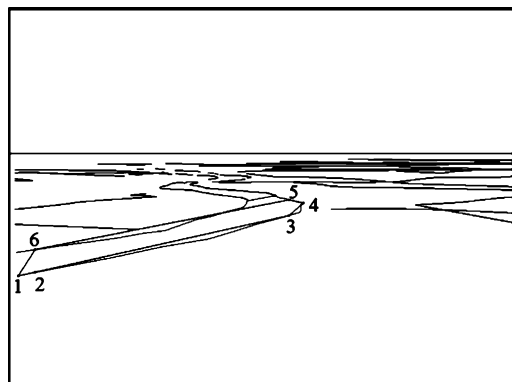
$$\cos \phi = \frac{(\bar{a}_1, \bar{a}'_1)}{|\bar{a}_1| \cdot |\bar{a}'_1|}$$
 угла между ними.

Если  $\cos \varphi > c$ , где, например,  $c = 0,966$ , что соответствует углу в 15 градусов, то точки  $M_i$  и  $M_{j_0}$  признаются «идентичными». В противном случае выносится решение о том, что точка, «идентичная» точке  $M_i$ , на втором контуре не найдена.

На рис. 3 приведены многоугольники, аппроксимирующие контуры реки соответственно на реальном и синтезированном изображениях. Каждой точке виртуального изображения ставится в соответствие некоторая точка реального изображения, эти соответствия и значения коэффициента  $\delta_{ij}$  приведены в табл. 1.



а)



б)

Рис. 3. Многоугольники, аппроксимирующие контуры реки на реальном и синтезированном изображениях

Таблица 1. Соответствия точек контуров и коэффициенты похожести

Точка ВИ	Точка РИ	Коэфф. $\delta$
1	1	0,49
2	2	0,51
3	3	0,42
4	4	0,56
5	7	0,45
6	8	0,73

В табл. 1 использованы следующие сокращения: ВИ – виртуальное изображение, РИ – реальное изображение.

Выбор ключевой точки из ограниченного множества пар соответствующих точек осуществляется по критерию  $j_0 = \arg \min_{j=1, n} \delta_{ij}$ . Однако соответствующие пары точек могут располагаться слишком близко к линии горизонта и к краям изображения, где оценки координат точек могут содержать наибольшие ошибки. Поэтому выбор ключевой точки дополняется сравнением на множестве пар точек со значением критерия  $\delta_{ij}$  близким к минимальному  $\delta_{ij_0}$  по положению этих точек относительно линии горизонта и краев изображения.

На основании анализа значений коэффициента  $\delta_{ij}$  в табл. 1 и дополнительных критериев в качестве ключевой точки для осуществления дальнейших процедур совмещения выбрана пара точек с номерами 4-4. Значение коэффициента  $\delta$  в этой точке больше чем, например, в точках 1, 2, 3, 5, но данная точка проходит по дополнительным условиям. Так, например, пары точек 1-1 и 2-2 находятся слишком близко к краю совмещаемых изображений. Пары 3-3 и 5-7 имеют худшие по сравнению с парой 4-4 значения дополнительного показателя по углам между векторами  $\vec{d}_1 = \vec{a}_1 + \vec{a}_2$  и  $\vec{d}_2 = \vec{b}_1 + \vec{b}_2$ , т. е. по ориентации углов  $\alpha_i = \angle M_{i-1}M_iM_{i+1}$  и  $\beta_j = \angle M'_{j-1}M'_jM'_{j+1}$  относительно координатных осей.

Следующие этапы алгоритма предсовмещения контуров связаны непосредственно с процедурой совмещения. Устранение несовпадения синтезированного и реального изображения целесообразно начать с определения угла между линиями горизонта на этих изображениях. Угол, на который нужно повернуть одно из двух изображений, определяется как угол между прямыми, аппроксимирующими линии горизонта на реальном и синтезированном изображениях. Линия горизонта на виртуальном изображении всегда четко выражена. Её положение рассчитывается в рамках специального комплекса программ, обеспечивающего формирование изображения, синтезированного по цифровой карте местности. На реальном изображении за линию горизонта принимается «линейный» массив пикселей максимальной длины в верхней части изображения не содержащий разрывов, за исключением, быть может, одиночных разрывов не более чем в один пиксель.

Оценки угловых коэффициентов  $b$ ,  $b_1$  уравнений прямых  $\tilde{\Gamma}: x = a + by$ ,  $\tilde{\Gamma}: x = a_1 + b_1y$ , аппроксимирующих линии горизонта соответственно на реальном и синтезированном изображениях можно найти стандартным методом наименьших квадратов (МНК) по формуле

Оценки угловых коэффициентов  $b$ ,  $b_1$  уравнений прямых  $\tilde{\Gamma}: x = a + by$ ,  $\tilde{\Gamma}: x = a_1 + b_1y$ , аппроксимирующих линии горизонта соответственно на реальном и синтезированном изображениях можно найти стандартным методом наименьших квадратов (МНК) по формуле

$$\hat{b} = \frac{n \cdot \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \cdot \sum_{i=1}^n y_i}{n \cdot \sum_{i=1}^n x_i^2 - \left( \sum_{i=1}^n x_i \right)^2} \quad (3)$$

Однако в вычислительном отношении эти оценки проще вычислить с помощью скалярного произведения

$$\hat{b} = \bar{q} \cdot \bar{x} \quad (4)$$

В формуле  $\bar{q} = \frac{3}{k(k+1)(2k+1)} (-k, -k+1, \dots, -1, 0, 1, \dots, k-1, k)$  – весовые коэффициенты оператора, обеспечивающего вычисление оптимальной МНК-оценки частной производной  $\frac{\partial f}{\partial x}$  [7];  $\bar{x} = (2k+1)$  – мерный вектор

тор ( $n = 2k + 1$ ) ординат линии горизонта. Применение формулы (4) обеспечивает выигрыш во времени вычисления оптимальной оценки углового коэффициента в среднем в 4 раза по сравнению с формулой (3).

После того как оценки  $\hat{b}$ ,  $\hat{b}_1$  угловых коэффициентов найдены вычисляем угол, на который нужно повернуть синтезированное изображение для его совмещения с реальным изображением по линиям горизонта.

$$\gamma = \arccos \frac{1 + \hat{b} \cdot \hat{b}_1}{\sqrt{1 + \hat{b}^2} \cdot \sqrt{1 + \hat{b}_1^2}} \quad (5)$$

Для изображений, приведенных на рис. 1, оценка угла (5) между линиями горизонта оказалась равной  $1,088^\circ$ . Это означает, что в навигационных данных, по которым построено синтезированное изображение, содержится ошибка  $\Delta\gamma = 1,088^\circ$  по углу крена. На рис. 4 приведен результат наложения исходного (рис. 1) синтезированного изображения, предварительно повернутого на угол  $1,088^\circ$ , на реальное изображение.

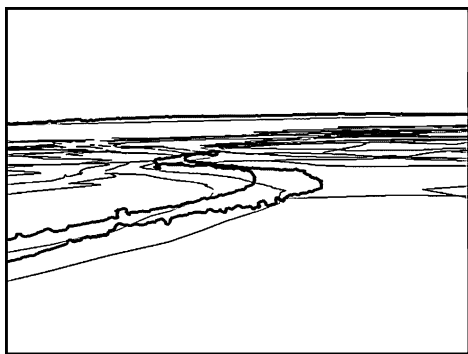


Рис. 4. Результат наложения реального и синтезированного изображений, после корректировки по углу крена

Можно отметить, что на визуальном уровне линии горизонта теперь совпадают. Однако одно изображение смещено относительно другого и, возможно, повернуто на некоторый угол. Для совмещения изображений с помощью операции сдвига вдоль вектора, выделим из найденного на первых двух этапах множества пар угловых точек одну *ключевую* точку. Выбор ключевой точки подчиним двум основным требованиям: она, во-первых, должна находиться в окрестности средней линии между нижним срезом экрана и линией горизонта и, во-вторых, должна быть *наиболее значимой* – принадлежать углу с максимально длинными сторонами.

Вычислив параметры сдвига  $\Delta x = x_M - x_{M'}$ ,  $\Delta y = y_M - y_{M'}$ , получим в нашем случае, что  $\Delta x = 68$ ,  $\Delta y = 17$  пикселей. Сместим синтезированное изображение на эти величины вдоль соответствующих осей. Теперь ключевые точки совпадут. Но одновременно возникнет смещение по вертикали линии горизонта синтезированного изображения относительно реального (рис. 5). Обозначим эту величину  $\Delta z$ .

Смещение линий горизонта обусловлено наличием проекционных искажений. Для учета этих искажений нужно подобрать такие значения смещений по высоте  $\Delta h$  и по углу тангажа  $\Delta\theta$ , которые привели бы к совмещению как линий горизонта, так и ключевых точек.

«Выбрать» это недосовмещение можно в рамках многошаговой процедуры за счет корректировки значений угла тангажа и высоты.

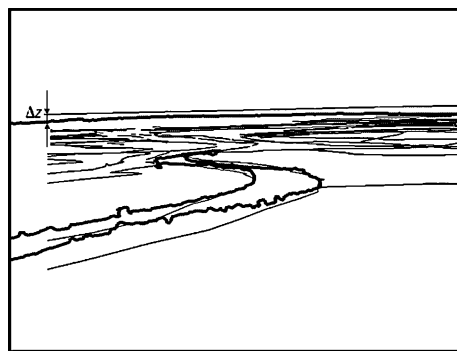


Рис. 5. Результат совмещения изображений по ключевым точкам

Корректировка по углу тангажа  $\theta$  заключается в следующем: на предыдущих стадиях выполнения алгоритма методом наименьших квадратов были найдены уравнения прямых, аппроксимирующих линии горизонта на реальном и синтезированном изображениях. По коэффициентам  $a$  и  $a_1$  этих прямых определяется, какой знак имеет поправка к углу тангажа, а также расстояние между этими прямыми в пикселях. В рамках итерационной процедуры производится уточнение истинного значения угла тангажа (критерием «уточнения» служит сокращение расстояния между линиями горизонта). Для этого задается некоторый шаг, с которым начинает изменяться угол тангажа. Для каждого значения этого угла и фиксированных значений остальных параметров генерируется новое синтезированное изображение и производится сравнение расстояний между прямыми, аппроксимирующими линии горизонта. Выполнение процедуры прекращается, когда расстояние между прямыми становится меньше некоторого заранее заданного порога (в рамках проводимых исследований – 2 пикселя).

Определение смещения  $\Delta h$  по высоте также носит итерационный характер. Выбирается некоторый шаг по данной координате и производится генерация ракурсов с различными значениями по высоте. Далее для каждой пары реального и синтезированного изображений рассчитывается некоторая мера близости изображений (в проводимом исследовании выбрана парная критериальная функция Рао [2]). В конечном итоге выбирается пара максимально близких по выбранному критерию изображений. Полученная пара из реального и синтезированного изображений и принимается за итоговый результат предсовмещения. Финальный результат цепочки таких преобразований приведен на рис. 6.

Итоговые значения корректировок угла тангажа и высоты таковы:  $\Delta\theta = -0,3^\circ$ ,  $\Delta h = -20$  м. Приведем значения корректировок по всем параметрам, задействованным в преобразованиях:  $\Delta x = 68$ ,  $\Delta y = 17$ ,  $\Delta\gamma = 1,088^\circ$ ,  $\Delta\theta = -0,3^\circ$ ,  $\Delta h = -20$  м.

Смещения  $\Delta x$  и  $\Delta y$  приведены в пикселях.

Для сравнения на рис. 7 приведен результат совмещения реального и синтезированного изображений корреляционно-экстремальным методом с использованием парной критериальной функции Рао [2]. Поиск глобаль-



ного экстремума критериальной функции проводился в пространстве 6 параметров  $(\Delta x, \Delta y, \Delta h, \Delta \psi, \Delta \theta, \Delta \gamma)$  на сетке, содержащей около  $10^6$  узлов.

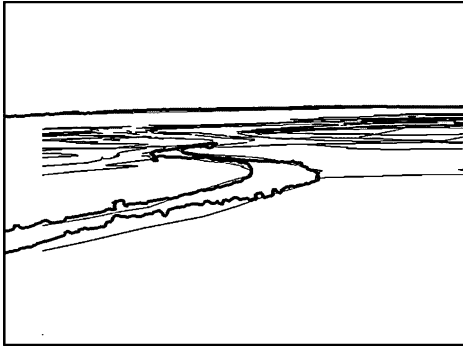


Рис. 6. Конечный результат совмещения изображений

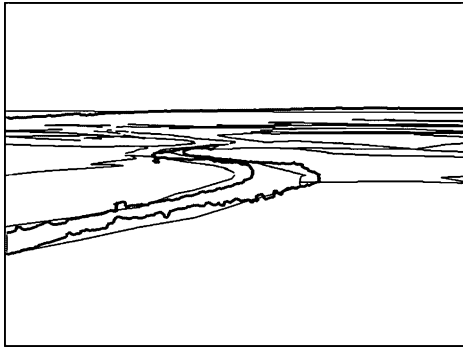


Рис. 7. Результат совмещения изображений корреляционным методом

Визуальное сопоставление результатов совмещения по ключевой точке (рис. 6) и корреляционным методом (рис. 7) позволяет говорить о том, что контуры реки совмещены в обоих случаях приблизительно одинаково. Однако по линиям горизонта изображения на рис. 7 недосовмещены. Для более точного сопоставления результатов совмещения изображений нужны критерии качества совмещения, основанные на количественных характеристиках.

Итоговые значения корректировок по каждому из шести параметров, полученные при корреляционном совмещении методом полного перебора таковы:  $\Delta x = 32\text{ м}$ ,  $\Delta y = 0$ ,  $\Delta h = -16\text{ м}$ ,  $\Delta \psi = -1^\circ$ ,  $\Delta \gamma = -0,2^\circ$ ,  $\Delta \theta = -0,2^\circ$ . Оценки смещений  $\Delta h$  и  $\Delta \theta$ , полученные методом ключевой точки близки к оценкам, найденным корреляционным методом; Смещения  $\Delta x$  и  $\Delta y$  трудно сопоставлять по двум причинам. Во-первых, они измеряются в разных единицах. Во-вторых, в методе ключевой точки не производятся коррекция по курсу. Поэтому оценки смещений  $\Delta x$  и  $\Delta y$  в методе ключевой точки содержат в своем составе и ошибку оценивания курса. Существенное несовпадение оценок ошибок по крену ( $\Delta \gamma = 1,088^\circ$  и  $\Delta \gamma = -0,2^\circ$ ) объясняется тем, что в методе ключевой точки линии горизонта совмещены качественно, а в корреляционном методе они недосовмещены.

**Оценка качества совмещения изображений**

От качества совмещения реального и синтезированного изображений зависит точность пересчета параметров пространственного и углового позиционирования

сенсоров, а также адекватность и эргономичность восприятия комбинированного изображения экипажем. Поэтому важно количественно оценить степень (качество) совмещения изображений для принятия решения о возможности использования результатов совмещения изображений для КЭНС и визуализации экипажу.

Качество совмещения линии горизонта можно оценить с двух позиций: есть ли систематическое смещение (рис. 8а) по одной из координат на плоскости и недосовмещение по углу (рис. 8б). Так, например в процессе совмещения реального и виртуального изображений после корректировки по углу крена смещение  $\Delta z$  линии горизонта по вертикали (рис. 5) синтезированного изображения относительно реального составило 8 пикселей.

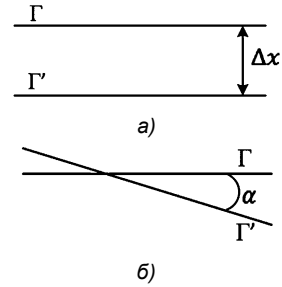


Рис. 8. Варианты недосовмещения линий горизонта

Для замкнутых областей  $D$  и  $D'$  можно ввести показатели качества совмещения  $\delta$  и  $\delta'$  – отношение площади пересечения замкнутых совмещаемых областей  $D \cap D'$  к площади соответственно первой и второй областей:

$$\delta = \frac{S_{D \cap D'}}{S_D} \cdot 100\%,$$

$$\delta' = \frac{S_{D \cap D'}}{S_{D'}} \cdot 100\%.$$

Оценка качества совмещения по первой формуле в составе (6) реального и синтезированного изображений (рис. 6) методом ключевой точки дала следующий результат:  $\delta = 73,18\%$ . Не очень высокое значение показателя качества совмещения объясняется структурой реального и синтезированного изображений. В нижней части изображения (рис. 3) контуры реки на реальном и синтезированном изображениях имеют существенные отличия, не позволяющие повысить качество совмещения.

Оценка качества совмещения по формуле (6) методом ключевой точки на каждом шаге этого метода приведена в табл. 2.

Оценка качества совмещения тех же изображений корреляционным методом с полным перебором узлов сетки (рис. 7) по первой формуле в составе (6) приводит к следующему результату:  $\delta = 72,4\%$ .

Близкие значения коэффициента  $\delta$  в методе ключевой точки и в корреляционном методе, основанном на полном переборе узлов сетки, позволяют говорить о возможности предсовмещения методом ключевой точки.

Отметим, что описанная методика оценки качества совмещения изображений может применяться без участия человека-оператора, что позволяет рассматривать её как возможное решение проблемы определения степени совмещения изображений на борту ЛА.

Таблица 2. Результаты поэтапной оценки качества совмещения изображений

Наименование этапа	Площадь области $D$ (количество пикселей)	Площадь области $D'$ (количество пикселей)	Площадь пересечения областей $D$ и $D'$	Значение коэффициента $\delta$
Исходные изображения (рис. 1)	14812	26567	7018	47,38%
Изображения после корректировки угла крена (рис. 4)	14812	26606	5438	36,71%
Изображения после сдвига вдоль осей (рис. 5)	14812	26606	8035	54,24%
Изображения после корректировок по высоте и углу тангажа (рис.6)	14812	23408	<b>10839</b>	<b>73,18%</b>

## Заключение

Жесткие ограничения на время обработки реального и синтезированного изображений на борту ЛА вынуждают искать быстрые алгоритмы решения каждой задачи в составе комплекса задач, решаемого бортовым вычислителем. Задача совмещения изображений с одновременным определением ошибок навигационных параметров ЛА, является в этом комплексе наиболее важной и сложной в вычислительном отношении.

Предложенный метод ключевой точки преследует цель сокращения вычислительных затрат на решение этой задачи. Он обеспечивает предварительное совмещение изображений. Оно, как это следует из данного исследования, может быть достаточно хорошим. Однако метод ключевой точки не решает в полном объеме задачу оценивания ошибок в навигационных параметрах.

Метод ключевой точки позволяет найти оценку  $\Delta y$  смещения по крену, близкие к истинным оценки смещений по высоте  $\Delta h$  и по углу тангажа  $\Delta \theta$ . Менее точными являются оценки смещений  $\Delta x$  и  $\Delta y$ . Полученная данным методом оценка 6-мерного вектора навигационных параметров  $(\hat{\Delta x}, \hat{\Delta y}, \hat{\Delta h}, 0, \hat{\Delta \theta}, \hat{\Delta \gamma})$  может использоваться в качестве исходного (нулевого) приближения для поиска «истинных» значений этого вектора корреляционным методом. Поиск в этом случае нужно будет осуществлять не по всему параллелепипеду в 6-мерном пространстве, а лишь в некоторой, достаточно малой, окрестности найденной точки и, возможно, лишь по трем параметрам  $\Delta x$ ,  $\Delta y$  и  $\Delta \psi$ . Это может позволить сократить количество вычислений парной критериальной функции с  $10^6$  значений при полном переборе до  $10^2$ .

## Литература

1. Баклицкий В.К., Бочкарев А.М. Методы фильтрации сигналов в корреляционно-экстремальных системах навигации. М.: Радио и связь, 1986. 1072 с.
2. Елесина С.И., Ефимов А.И. Отбор критериальных функций для систем улучшенного и комбинированного видения// Известия ТулГУ, технические науки, вып. 9, ч.1. 2013. - С. 229-236.
3. Новиков А.И., Саблина В.А., Горячев Е.О. Приме-

нение контурного анализа для совмещения изображений// Известия ТулГУ, технические науки, вып. 9, ч.1. 2013. – С. 260-270.

4. Novikov A.I., Sablina V.A., Nikiforov M.B., Loginov A.A. Contour Analysis and Image superimposition Task in Computer vision Systems//11th International Conference on Pattern Recognition and Image Analysis: New Information Technologies (PRIA– 11–2013). Samara, 2013. vol. 1. – PP. 282-285.

5. Dmitry Chetverikov and Zsolt Szabo. A Simple and Efficient Algorithm for Detection of High Curvature Points in Planar Curves, Proc. 23rd Workshop of the Austrian Pattern Recognition Group, 1999, pp. 175-184.

6. Непомнящий П.В., Юрин Д.В. Поиск опорных точек на векторных изображениях путем детектирования угловых структур с помощью статистической оценки гипотез. // Graphicon 2002 proceedings, <http://www.graphicon.ru/2002/>.

7. Новиков А.И. Алгоритмы выделения границ полезных сигналов // Вестник РГПУ. №2 (выпуск 24). – Рязань. - 2008. - С. 11-15.

## PRE-COMBINING IMAGES AND QUALITY EVALUATION METHODS OF COMBINING

*Novikov A.I., Efimov A.I.*

The report examines the problems of development of enhanced and combined vision systems of aircraft. It focuses on the issue of different spectral image combining. The authors propose a methodology of images pre-combining based on finding a «key point» for binding of combined images. Authors describe the steps for correcting the coordinates of the aircraft (roll and pitch angles and height) which significantly improves the quality of combining images obtained from a camera on the board of aircraft and images generated by digital area map. In the report demonstration of new algorithm of pre-combining images is made. It also shows high quality of transformation series over the virtual image and result of images combining. Thus it is possible to obtain estimates of navigational error of the aircraft. Authors demonstrated the possibility of quality rating not only visually, but also by using quantitative index of quality.

УДК 004.932

## ОБРАБОТКА ОБЛАСТЕЙ НАЛОЖЕНИЙ ОБЪЕКТОВ ПРИ ПОСТРОЕНИИ КАРТ ГЛУБИНЫ ДЛЯ ВИДЕО

*Матюнин С.Б., аспирант факультета вычислительной математики и кибернетики МГУ  
им. М. В. Ломоносова, e-mail: smatyunin@graphics.cs.msu.ru;*

*Ватолин Д.С., к. ф.-м. н., с. н. с. факультета вычислительной математики и кибернетики МГУ  
им. М. В. Ломоносова, e-mail: dmitriy@graphics.cs.msu.ru.*

**Ключевые слова:** карта глубины, видео, обработка изображений, оптический поток, сжатие стереовидео.

### Введение

В данной статье рассматривается задача полуавтоматического построения карты глубины для видео. Под термином «карта глубины» для некоторого кадра будем понимать изображение, содержащее вместо яркостной или цветовой информации расстояния до соответствующих видимых точек сцены. Карты глубины, полученные с помощью прямого измерения специализированными приборами [1] или посредством анализа стереопары [2], применяются для трехмерного моделирования. В киноиндустрии карты глубины применяются как промежуточный этап преобразования обычного кино (будем его называть 2D кино или 2D видео) в стереоформат. Процесс создания карт глубины для видео подразумевает создание карты расстояний для каждого кадра. Данная задача в общем случае не имеет решения. На практике применяются ручная разметка и различные эвристические методы [3, 4, 5]. При ручной обработке трудозатраты можно уменьшить с помощью компьютерных инструментов. Это могут быть как инструменты обработки изображений общего назначения: кисти, заливка, сегментация и т.п., а также специализированные: фильтры карты глубины, алгоритмы распространения ручной разметки, методы автоматического построения карты глубины для видео

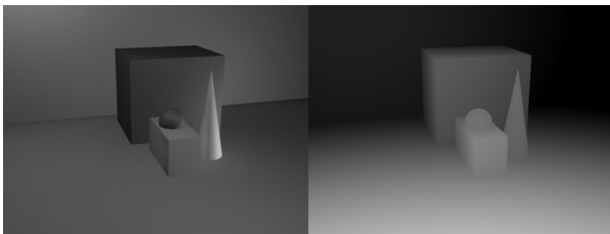


Рис. 1. Пример трехмерной сцены (слева) и карты глубины для нее (справа)

Один из распространенных полуавтоматических подходов [6, 7, 8, 9] состоит в том, что пользователь вручную создает карту глубины для некоторых кадров видео. Будем называть эти кадры ключевыми. Объекты между соседними кадрами, как правило, не сильно изменяют свою позицию, а следовательно, и глубину. Поэтому карту глубины для остальных кадров можно стро-

*Рассматривается задача полуавтоматического построения карты глубины для видео. Предложен метод обработки областей наложений объектов для повышения качества карт глубины. Информация о таких областях аккумулируется при последовательной обработке кадров видео, а затем используется для маскирования ненадежных интерполированных областей. Приведено сравнение алгоритмов поиска областей наложений объектов и способов их использования в рассматриваемой задаче.*

ить автоматически – интерполируя ее для необработанных кадров. Для повышения качества интерполированной карты глубины предлагается отдельно обрабатывать области наложения объектов в видео.

Пример трехмерной сцены (слева) и карты глубины для нее (справа) представлен на рис. 1. Яркость пикселей карты глубины соответствует расстоянию от точки наблюдения до поверхности объекта. Большим значениям яркости соответствует более близкие поверхности.

### Обзор методов построения карты глубины

В работах [6, 8, 10] предлагаются методы полуавтоматической разметки карты глубины для видео, основанные на использовании пользовательского ввода для некоторых кадров последовательности (ключевых кадров). Исходная разметка распространяется с ключевых кадров на остальные с учетом информации о движении, цвете и других характеристиках, которые можно извлечь из видео.

В [10] для интерполяции карты глубины между ключевыми кадрами составляется система уравнений, ограничивающих разность по глубине между соседними пикселями в кадре и соответствующими пикселями в соседних кадрах. Системой также задается глубина для размеченных пикселей и глубина пикселей, найденная с помощью классификатора по цвету. Полученная система решается методом наименьших квадратов. Решение соответствует карте глубины для всей последовательности. Такой подход требует решения системы уравнений большой размерности. Это приводит к необходимости уменьшать размерность за счет разрешения искомой карты глубины. Аналогичным образом карта глубины строится в работе [11]. Для каждого пикселя видео-последовательности задается уравнение вида

$$d(p) = \sum_{q \in N(p)} \omega_{pq} d(q), \quad (1)$$

где  $d(p)$  – глубина пикселя  $p$ ,  $N(p)$  – множество пик-

селей, состоящее из 8 соседних пикселей на текущем кадре и соответствующего соседнего пикселя с предыдущего кадра,  $\omega_{pq}$  – вес соседнего пикселя, обратно пропорциональный разности цветов пикселей  $p$  и  $q$ . Соседний пиксель с предыдущего кадра определяется с учетом векторов оптического потока. Из полученных уравнений составляется линейная система  $Ax = 0$ . Для нахождения карты глубины решается задача квадратичного программирования:  $\min_x \|Ax\|^2$  при ограничениях  $Ex \geq f$ , заданных пользовательским вводом.

Чаще используется подход, который позволяет последовательно вычислять карту глубины для соседнего кадра. В работе [8] глубина для каждого следующего кадра вычисляется с помощью билатерального фильтра, веса которого зависят от геометрического и фотометрического расстояния между пикселями. Возможные ошибки исправляются с помощью компенсации движения. За счет применения компенсации движения только на этапе постобработки данный подход не применим для видеопоследовательностей с сильным движением.

В работе [12] межкадровая интерполяция используется для увеличения частоты кадров карты глубины. Интерполяция осуществляется за счет использования векторов движения, полученных по соответствующему 2D-видеопотоку.

В задаче сжатия карты глубины также применяются аналогичные подходы. Например, в методе [11] компенсация движения используется при предсказании карты глубины для очередного кодируемого кадра.

Подход, аналогичный описанному в [8, 12, 13, 14], мы использовали как базовый для построения карты глубины. Он позволяет последовательно получать карту глубины для каждого следующего кадра. Размерность решаемой задачи при этом значительно меньше, чем при построения карты глубины для всей последовательности сразу.

### Построение областей открытия для пары кадров

При обработке сцен, содержащих движущиеся объекты, обычно применяется оценка движения. При поиске векторов движения критичной является корректная обработка областей открытия, то есть тех областей, которые были не видны на одном кадре, но на другом из-за движения объектов стали видимы. В данной работе мы используем существующие методы построения областей открытия для пары кадров. На их основе предлагается строить маску «доверия» для интерполированной карты глубины, а далее использовать ее для улучшения качества результата.

Задача построения областей открытия по видео в общем случае не имеет точного решения. Например, для двух последовательных кадров можно определить области открытия несколькими способами, в зависимости от реального направления движения (см. рис. 2). Поэтому на практике используются различные эвристические подходы.

Один из наиболее часто используемых подходов [17] основан на анализе ошибки компенсации движения между двумя кадрами. Область кадра, где ошибка компен-

сации высока в обоих направлениях, рассматривается как область наложения. В силу простоты критерия данный подход имеет высокую скорость работы, однако он неустойчив к шуму, дает большое число ложных срабатываний при малых значениях порога  $\Theta$  и наоборот, малое число верных срабатываний при больших  $\Theta$ . Аналогичный подход (LRC) основан на анализе разности векторов движения от одного кадра к следующему и обратно. В работе [17] также предложен геометрический подход, основанный на анализе плотности концов векторов движения на опорном кадре. Так как на кадре  $t$  не существуют соответствия для точек из области закрытия кадра  $t-1$ , поэтому в таких областях функция плотности будет иметь малые значения. Данный подход обладает хорошей устойчивостью, так как опирается на надежные вектора движения в областях закрытия вместо некорректно определенных векторов в областях открытия.

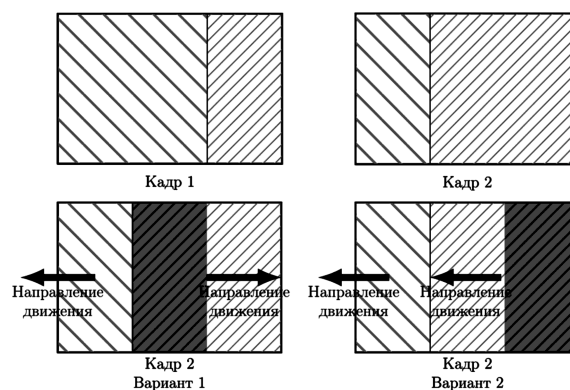


Рис. 2. Неоднозначность определения области открытия по двум кадрам. Возможная область открытия отмечена серым цветом, объекты отмечены штриховкой

### Обозначения

Пусть размер кадра равен  $W \times H$ . Обозначим через  $I_t(p)$  яркость в точке с координатами  $p = (p_x, p_y)$ ,  $p \in [0, W-1] \times [0, H-1]$  на кадре в момент времени  $t$ . При этом запись  $I_t$  будет обозначать весь кадр в момент времени  $t$ . Оценка движения – поиск соответствия между двумя кадрами, один из которых называется базовым, а другой – опорным. Для каждой точки  $x_{t_0}$  базового кадра  $I_{t_0}$  алгоритм оценки движения находит соответствующую ей точку  $x_{t_1}$  опорного кадра  $I_{t_1}$ . Разность между координатами этих точек называется вектором движения  $V_{t_0, t_1}(x_{t_0})$ , вычисленным в точке  $x_{t_0}$  базового кадра  $I_{t_0}$ :  $V_{t_0, t_1}(x_{t_0}) = x_{t_1} - x_{t_0}$ . Совокупность  $V_{t_0, t_1}$  всех векторов движения  $V_{t_0, t_1}(x_{t_0})$  базового кадра называется полем векторов движения для кадров  $I_{t_0}$  и  $I_{t_1}$ .

При движении объектов на видео возникают области их взаимного наложения. Можно выделить два типа наложений на заданной паре кадров [15]:

1. **Открытие** – область базового кадра, отсутствующая на опорном кадре;
2. **Закрытие** – область опорного кадра, отсутствующая на базовом кадре.

Компенсация движения – процесс построения скомпенсированных кадров  $\tilde{I}_t$  на основе опорного кадра  $I_{t_0}$  с использованием информации о движении между исходными кадрами  $I_{t_0}$  и  $I_t$ :

$$\tilde{I}_t(x) = I_{t_0}(x + V_{t_0,t_1}(x_{t_0})). \quad (2)$$

Для распространения карты глубины  $D$  с кадра  $t-1$  на кадр  $t$  используются векторы движения  $V_{t-1,t}$ , найденные по исходному 2D видео с помощью преобразования компенсации движения  $MC$ :

$$MC_{t,t-1} * D_{t-1}(x) \stackrel{def}{=} D_t(x + V_{t-1,t}(x)). \quad (3)$$

Если известна карта глубины  $D_{t-1}$  для предыдущего кадра, то можно экстраполировать ее для кадра  $t$ :

$$\hat{D}_t(x) = MC_{t,t-1} * D_{t-1}(x). \quad (4)$$

Для краткости обозначим преобразование для всего кадра  $\hat{D}_t = MC_{t,t-1} * D_{t-1}$ . Если значение  $x + V_{t-1,t}(x)$  выходит за допустимые пределы, оно приводится к ближайшему допустимому. Преобразование можно применять последовательно, начиная с некоторого заданного кадра  $k$ , распространяя карту глубины на всю видео-последовательность:

$$\hat{D}_t = MC_{t,t-1} * MC_{t-1,t-2} * \dots * MC_{k+1,k} * D_k. \quad (5)$$

Положим  $\hat{D}_k \stackrel{def}{=} D_k$ . При таком подходе качество полученной карты глубины напрямую зависит от алгоритма оценки движения. Основным недостатком такого подхода является то, что в областях открытия, которые присутствуют на базовом кадре и отсутствуют на опорном, векторы движения не определены. Большинство алгоритмов оценки движения опирается на скомпенсированную межкадровую разность и поэтому в областях открытия они дают некоторые неадекватные (зависящие от алгоритма оценки движения и входного изображения) значения векторов движения.

В областях открытия векторы движения определяются некорректно, поэтому «копирование» карты глубины для области фона может происходить из области объекта (рис. 3).

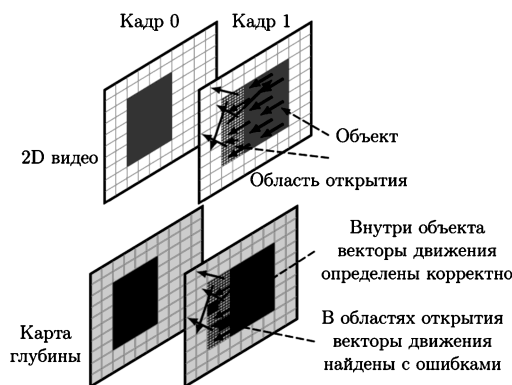


Рис. 3. Причина возникновения ошибки интерполяции карты глубины при использовании компенсации движения.

### Обработка областей открытия

Компенсация карты глубины (4) в областях открытия приводит тому, что за движущимся объектом образуется «шлейф» неправильной глубины (см. рис. 4). Перед объектом карта глубины получена с использованием более надежных векторов движения, поэтому результат в этой области корректный. При распространении с разных направлений корректные значения карты глубины получаются по разные стороны от объекта. Это можно использовать для улучшения результата обработки.

Рассмотрим последовательность кадров длины  $n$  (сцену длины  $n$ ):  $I_0, \dots, I_{n-1}$ . Допустим, что известна карта глубины  $D_0$  и  $D_{n-1}$  для первого и последнего кадров соответственно (рис. 4). Рассмотрим объект движущийся поступательно слева направо. При распространении от 0-го кадра до некоторого кадра  $k \in [0, n-1]$  «шлейф» находится слева от объекта (рис. 4 е), а при распространении от последнего кадра – справа от объекта (рис. 4д).

Результат работы алгоритма может быть улучшен, путем слияния результатов распространения с нескольких направлений. Предположим, имеется  $m$  размеченных ключевых кадров с индексами  $k_0, \dots, k_{m-1}$ . Для каждого кадра последовательности можно построить  $m$  вариантов экстраполированной карты глубины  $\hat{D}^0, \dots, \hat{D}^{m-1}$ . Можно использовать простое усреднение результатов распространения с разных направлений:

$$\hat{D}_t^{avg}(x) = \frac{1}{m} \sum_{i=0}^{m-1} \hat{D}_t^i(x) \quad (6)$$

где  $\hat{D}_t^{avg}$  – итоговая карта глубины,  $m$  – количество направлений распространения (для экспериментов использовалось распространение с первого и с последнего кадра, т.е.  $m = 2$ ). Такой подход не дает удовлетворительного результата, т. к. «шлейф» остается с обеих сторон объекта (рис. 8 а, б). Другой подход состоит в том, что результаты с двух направлений смешиваются с весами, зависящими от расстояния до соответствующего ключевого кадра, например:

$$\hat{D}_t^w(x) = \frac{\sum_{i=0}^{m-1} w_t^i \hat{D}_t^i(x)}{\sum_{i=0}^{m-1} w_t^i} \quad (7)$$

$w_t^i = 1$ , при  $t = k_i$  и убывает до нуля при приближении  $t$  к соседнему ключевому кадру  $k_{i-1}$  или  $k_{i+1}$  (рис. 5).

Таким образом, для ключевого кадра берется только исходная карта глубины, а на промежуточных кадрах, расположенных между размеченными, используется среднее значение из результатов распространения с двух ближайших ключевых кадров. Когда область открытия образуется близко к ключевому кадру, веса ошибочных пикселей будут большими. Это увеличивает заметность ошибок распространения. Для более эффективного удаления шлейфа необходимо производить смешивание направлений с учетом положения областей открытия.

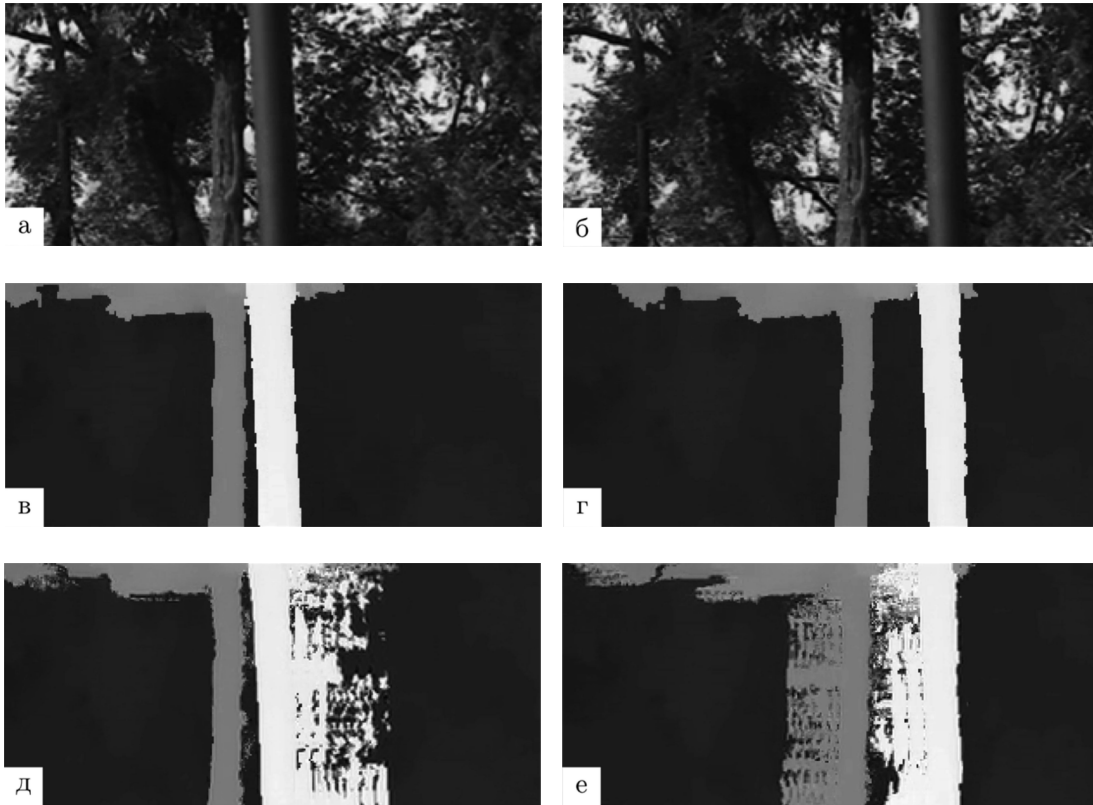


Рис. 4. Возникновение затеканий при покадровом распространении карты глубины при использовании базового подхода. Фрагмент видеопоследовательности «Road» [16]. а) Кадр №1; б) Кадр №7; в) Эталонная карта глубины для кадра №1; г) Эталонная карта глубины для кадра №7; д) Результат покадровой интерполяции с 7-го до 1-го кадра; е) Результат покадровой интерполяции с 1-го до 7-го кадра

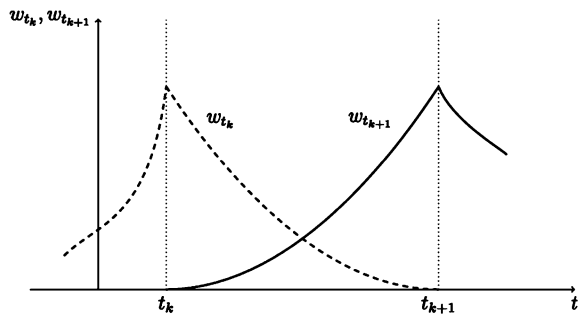


Рис. 5. Пример весовой функции для объединения результатов интерполяции для ключевых кадров  $t_{k-1}$  и  $t_{k+1}$

### Построение меры доверия

Покадровую маску областей открытия без дополнительной обработки трудно применить для улучшения качества интерполированной карты глубины. Одним из возможных подходов к заполнению областей открытий является синтез изображения в этих областях (inpainting, например [18]). Эта задача имеет множество решений для изображений. Для карт глубины она фактически представляет собой задачу построения карты глубины по одному изображению. Чтобы избежать синтеза карты глубины, и максимально использовать достаточно надежную информацию о глубине на ключевых кадрах и о движении объектов между кадрами, мы предлагаем использовать интерполированную карту глубины с нескольких ключевых кадров и построенные области открытия.

Мы предлагаем аккумулировать информацию об областях открытия при переходе от кадра к кадру. При этом области, помеченные как открытие на текущем кадре, в общем случае могут двигаться на последующих кадрах. Пусть  $\Omega(x)$  – карта областей открытия для кадра  $t$ :

$$\Omega(x) = \begin{cases} 1, & \text{если } x - \text{точка, не принадлежащая} \\ & \text{области открытия,} \\ 0, & \text{иначе.} \end{cases} \quad (8)$$

Тогда аккумулированная маска «доверия»  $\bar{\Omega}^j$  при проходе от ключевого кадра  $k_j$  будет строиться следующим образом:

$$\hat{\Omega}_{k_j}^j = \Omega_{k_j} \quad (9)$$

$$\hat{\Omega}_t^j(x) = \min(MC_{t,t-1} * \hat{\Omega}_{t-1}^{k_j}, \Omega_t) \quad (10)$$

Процесс построения маски областей открытия схематично показан на рис. 6. Если при компенсации движения координаты выходят за допустимую область, то значение  $MC_{t,t-1} * \hat{\Omega}_{t-1}^{k_j}$  принимается равным нулю, т.к. фактически это означает движения из-за границы кадра. Пиксель кадра  $t$  с координатой  $x_t$  помечается как область открытия в двух случаях: 1) если он лежит в области открытия на текущем кадре; 2) если соответствующий ему с учетом движения пиксель  $x_{t-1}$  на кадре  $t-1$  в аккумулированной маске  $\hat{\Omega}_{t-1}^j$  принадлежит области открытия.

### Использование меры доверия

Построенную на основе областей открытия маску доверия предлагается использовать для повышения качества итоговой карты глубины. Из предварительно построенных вариантов интерполяции карты глубины больший приоритет в каждом пикселе имеют варианты с большей мерой доверия. За счет этого уменьшается вклад в итоговый результат областей, в которых интерполяция производилась по некорректным векторам движения. Мы разработали и протестировали несколько способов использования построенной меры доверия.

Наиболее простым вариантом является выбор в каждой точке того направления распространения  $s_t^\Omega$ , которое соответствует большему значению меры доверия:

$$s_t^\Omega = \arg \max(\hat{\Omega}_t^i) \quad (11)$$

Результирующим значением глубины будет

$$D_t^\Omega = D_t^{s_t^\Omega} \quad (12)$$

При таком подходе результат для областей видео, не являющихся областями открытия, определяется порядком обхода  $\hat{\Omega}_t^i$ . Если  $D_t^i$  значительно отличаются для разных  $i$ , и на некотором кадре  $t$   $s_t^\Omega$  меняется для значительной части кадра, то результат будет содержать резкий переход по глубине. Обычно это отрицательно сказывается на качестве видео, обработанного с использованием карты глубины, например при переводе видео из 2D формата в 3D.

Чтобы уменьшить количество резких изменений глубины, можно в областях, где  $\hat{\Omega}_t^i = 1$ , применить дополнительный критерий, например, расстояние до ключевого кадра. Среди направлений распространения можно выбирать те, которые имеют максимальное значение меры доверия:

$$J_t^\Omega \left\{ j : \hat{\Omega}_t^j = \hat{\Omega}_t^{s_t^\Omega} \right\} \quad (13)$$

Из них выбираем направление, которое имеет наименьшее расстояние до ключевого кадра:

$$s_t^{\Omega Dist}(x) = \arg \min_{j \in J_t^\Omega} (|j - k_j|) \quad (14)$$

Найденное направление используется в качестве итогового для данной точки:

$$D_t^{\Omega Dist} = D_t^{s_t^{\Omega Dist}(x)} \quad (15)$$

При таком подходе резкий скачок может быть только на середине отрезка между двумя ключевыми кадрами, поэтому можно усреднить глубину для направлений с близкими к максимальному значениями меры доверия.

$$J_t^{\Omega T} \left\{ j : \left| \hat{\Omega}_t^j - \hat{\Omega}_t^{s_t^\Omega} \right| < T \right\} \quad (16)$$

где  $T$  – заданный порог. Итоговое значение получаем усреднением выбранных кандидатов:

$$D_t^{\Omega T} = \frac{1}{|J_t^{\Omega T}|} \sum_{j \in J_t^{\Omega T}} D_t^j \quad (17)$$

Усреднение может производиться с весами, убывающими при возрастании расстояния до ключевого кадра:

$$D_t^{\Omega \omega T}(x) = \frac{1}{\mu} \sum_{j \in J_t^{\Omega T}} \omega(|j - k_j|) D_t^j \quad (18)$$

Здесь  $\mu$  – нормировочный коэффициент. Более значимыми при близких значениях меры доверия считаются кандидаты, которые расположены ближе к соответствующим ключевым кадрам. Далее будет приведено сравнение описанных подходов.

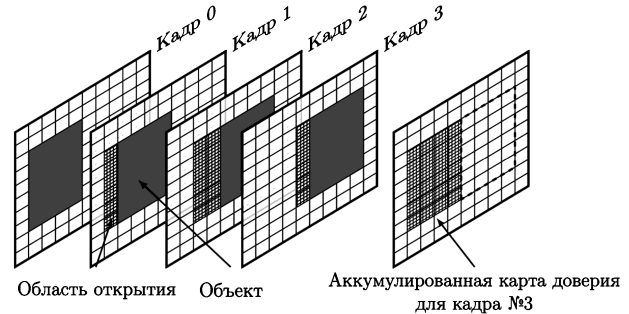


Рис. 6. Построение аккумулятивной карты областей открытия для нескольких кадров

### Результаты

Разработанный метод использования областей открытия был протестирован на наборе видеопоследовательностей с построенной картой глубины [16]. Для тестирования качества интерполяции карты глубины разработанному алгоритму на вход подавались карты глубины для двух ключевых кадров и отрезков видео между ними. Полученная карта глубины сравнивалась с эталонной. Для этого измерялось пиковое отношение сигнал/шум (PSNR). Результаты представлены на рис. 7. Для оценки поля векторов движения использовался алгоритм построения оптического потока [19].



Рис. 7. Сравнение методов построения областей открытия по качеству распространения карты глубины с учетом областей открытия

В качестве опорного метода было взято усреднение кандидатов без учета областей открытия ( $\hat{D}^{avg}$ ). Второй вариант алгоритма помечал как ненадежные те области, которые были скомпенсированы по движению из-за границ кадра. Это позволило повысить качество построенной карты глубины более чем на 10 дБ. Дальнейшее улучшение качества было достигнуто за счет обработки областей открытия внутри кадра. Наилучшие результаты по PSNR были достигнуты с использованием подхода

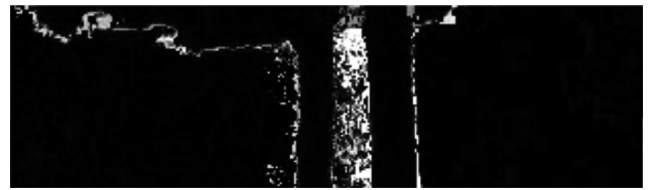
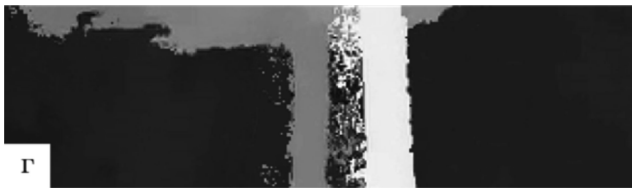
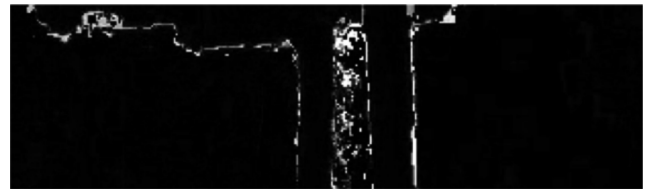
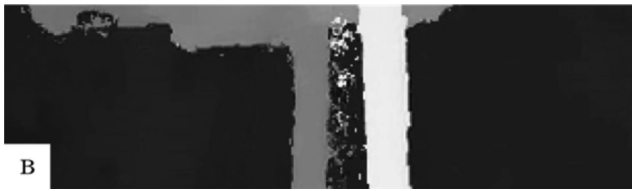
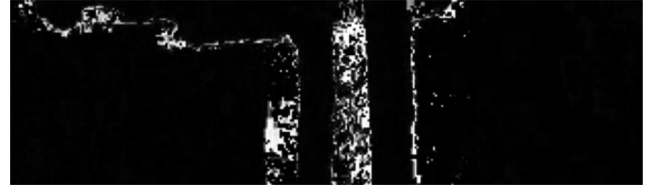
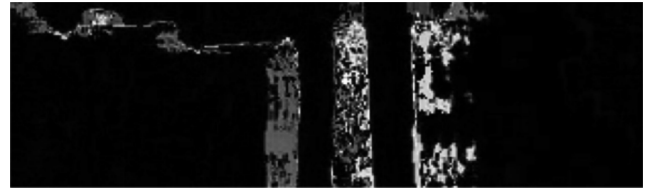
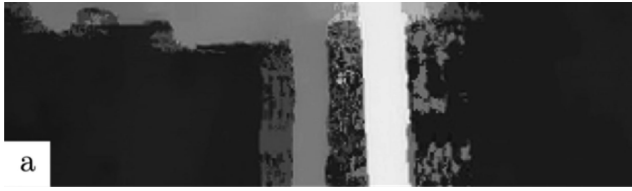


Рис. 8. Карта глубины, построенная с помощью различных методов обработки областей открытия (слева) и визуализация отличий по сравнению с эталоном (справа)

LRC, но этот метод порождает большое количество выбросов вблизи границ объектов (рис. 8 г). Геометрический подход имеет близкие показатели PSNR и продемонстрировал хорошую устойчивость вблизи границ (рис. 8 в). Среди протестированных подходов объединения результатов распространения карты глубины с нескольких кадров наилучшим оказался подход, использующий усреднение кандидатов с близкими к максимальному значениями меры доверия ( $D_t^{\Omega_{ot}}$ ).

Дополнительно было протестировано влияние обработки областей открытия на результаты работы алгоритма сжатия стереовидео. Использованный подход более подробно описан в [14]. Его суть в следующем. Стереовидео представляется в формате 2D+глубина. 2D видео кодируется с использованием классического кодека общего назначения. В данном случае использовался кодек x264. Для эффективного сжатия карты глубины используется ее согласованность с 2D видео, за счет чего можно повышать качество декодированной карты глубины. Карты глубины прореживаются как по времени, так и по пространственным измерениям и сжимаются. На этапе декодирования разжатое 2D видео используется для восстановления исходного разрешения и прореженных кадров карты глубины.

Декодированная карта глубины сравнивалась с исходной по пиковому соотношению сигнал/шум (PSNR). Для оценки влияния обработки областей открытий на алгоритм сжатия были построены RD-кривые. Исполь-

зование детектора областей открытия при декодировании карты глубины позволило повысить качество декодированной карты глубины до 2 дБ без изменения размера сжатых данных (рис. 9).

### Заключение

В данной работе был проведен сравнительный анализ трех методов определения областей открытия для видео в применении к задаче межкадровой интерполяции карты глубины. Наиболее перспективным оказался геометрический метод анализа векторов движения из работы [17]. Этот метод показал близкие к наилучшим значения PSNR, и в то же время показал лучшие результаты на границах объектов – меньше артефактов чем у метода LRC, что важно для карт глубины.

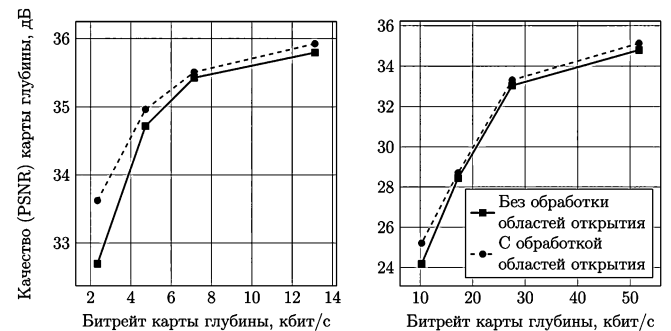


Рис. 9. Сравнение алгоритмов сжатия карты глубины на основе межкадрового распространения с обработкой областей открытия и без нее



Было предложено построение маски доверия для интерполированной карты глубины на основе пок кадровых областей открытия и несколько вариантов её использования. По результатам тестов наилучшим оказался метод с использованием расстояния до ключевого кадра со взвешенным усреднением. На тестовом наборе он показал улучшение показателя качества (PSNR) интерполированной карты глубины более чем на 1 дБ.

Обработка областей открытия позволила повысить качество карты глубины (до 2 дБ при низких значениях битрейта). Соотношение качество/битрейт регулировалось количеством использованных ключевых кадров карты глубины.

На визуализации (рис. 8) более светлые области соответствуют большей ошибке. Здесь: а) карта глубины, полученная простым усреднением; б) результат с обработкой областей открытия, полученных по ошибке компенсации движения; в) результат с обработкой областей открытия, полученных с помощью геометрического метода; г) результат с обработкой областей открытия, полученных по критерию LRC.

### Литература

1. Blais F. Review of 20 years of range sensor development. // *Journal of Electronic Imaging*. – 2004. – Vol. 13, no. 1. – Pp. 231–243.
2. Ogale A. S., Aloimonos Y. Shape and the Stereo Correspondence Problem // *International Journal of Computer Vision*. – 2005. – Vol. 65, no. 3. – Pp. 147–162.
3. Zhuo S., Sim T. On the Recovery of Depth from a Single Defocused Image // *Computer Analysis of Images and Patterns* / Ed. by Xiaoyi Jiang, Nicolai Petkov. – Springer Berlin Heidelberg, 2009. – Vol. 5702 of *Lecture Notes in Computer Science*. – Pp. 889–897.
4. Battiato S., Curti S., La Cascia M. et al. Depth map generation by image classification. – 2004. <http://dx.doi.org/10.1117/12.526634>.
5. Saxena A., Ng A., Chung S. Learning Depth from Single Monocular Images // *IEEE Neural Information Processing Systems*. – 2005. – Vol. 18.
6. Li Z., Cao X., Dai Q. A novel method for 2D-to-3D video conversion using bi-directional motion estimation // *2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. – 2012. – Pp. 1429–1432.
7. Rapid 2D-to-3D conversion / P. V. Harman, J. Flack, S. Fox, M. Dowley / Ed. by A. J. Woods, J. O. Merritt, S. A. Benton, M. T. Bolas. – Vol. 4660. – SPIE, 2002. – Pp. 78–86. <http://link.aip.org/link/?PSI/4660/78/1>.
8. Varekamp C., Barenbrug B. Improved depth propagation for 2D to 3D video conversion using key-frames // *IET Conference Publications*. – 2007. – Vol. 2007, no. CP534. – Pp. 29–29. <http://link.aip.org/link/abstract/IEECPS/v2007/iCP534/p29/s1>.
9. Practical temporal consistency for image-based graphics applications / M. Lang, O. Wang, T. Aydin et al. – Vol. 31. – New York, NY, USA: ACM, 2012. – Pp. 34:1–34:8. <http://doi.acm.org/http://doi.acm.org/10.1145/2185520.2185530>.
10. Guttman M., Wolf L., Cohen-Or D. Semi-automatic stereo extraction from video footage. // *ICCV'09*. – 2009. – Pp. 136–142.
11. Video stereolization: Combining motion analysis with user interaction / M. Liao, J. Gao, R. Yang, M. Gong // *Visualization and Computer Graphics, IEEE Transactions on*. – 2012. – Vol. 18, no. 7. – Pp. 1079–1088.
12. Choi J., Min D., Sohn K. 2D-plus-depth based resolution and frame-rate up-conversion technique for depth video // *Consumer Electronics, IEEE Transactions on*. – 2010. – November. – Vol. 56, no. 4. – Pp. 2489–2497.
13. De Silva D.V.S.X., Fernando W. A C, Yasakethu S. L P. Object based coding of the depth maps for 3D video coding // *Consumer Electronics, IEEE Transactions on*. – 2009. – August. – Vol. 55, no. 3. – Pp. 1699–1706.
14. Matyunin S., Vatolin D. 3D Video Compression Using Depth Map Propagation // *Multimedia Communications, Services and Security* / Ed. by A. Dziech, A. Czyżewski. – Springer Berlin Heidelberg, 2013. – Vol. 368 of *Communications in Computer and Information Science*. – Pp. 153–166.
15. Гришин С.В. Программная система для преобразования частоты кадров цифровых видео сигналов: Дис. канд. физ.-мат. наук: 05.13.11 / МГУ. – М., 2009.
16. Consistent depth maps recovery from a video sequence / Guofeng Zhang, Jiaya Jia, Tien-Tsin Wong, Hujun Bao // *Pattern Analysis and Machine Intelligence, IEEE Transactions on*. – 2009. – Vol. 31, no. 6. – Pp. 974–988.
17. Ince S., Konrad J. Geometry-based estimation of occlusions from video frame pairs // *Acoustics, Speech, and Signal Processing (ICASSP'05). IEEE International Conference on / IEEE*. – Vol. 2. – 2005. – Pp. ii–933.
18. Telea A. An image inpainting technique based on the fast marching method // *Journal of graphics tools*. – 2004. – Vol. 9, no. 1. – Pp. 23–34.
19. Ayvaci A., Raptis M., Soatto S. Sparse Occlusion Detection with Optical Flow // *International Journal of Computer Vision*. – 2012. – Vol. 97, no. 3. – Pp. 322–338. <http://dx.doi.org/10.1007/s11263-011-0490-7>.

## OCCLUSION PROCESSING FOR DEPTH MAP ESTIMATION

*Matyunin S., Vatolin D.*

We develop a system for semi-automatic converting 2D video into 3D stereoscopic content. Traditional pipeline of conversion includes mainly manual depth map mark-up for video and automatic views generation. Instead of manual depth map creation, we use only limited user input for several frames in video and automatically propagate it on other frames.

Most video processing algorithms fail in occlusion area where motion estimation is unreliable. We propose a technique for occlusion processing. We evaluated several occlusion area detection and usage strategies. The proposed method was applied to depth map creation and depth map compression.

## АЛГОРИТМЫ ИДЕНТИФИКАЦИИ ГИПЕРСПЕКТРАЛЬНЫХ ХАРАКТЕРИСТИК ОБЪЕКТОВ В ЗАДАЧАХ ДИСТАНЦИОННОГО ЗОНДИРОВАНИЯ ЗЕМЛИ

*Демидова Л.А., профессор РГРТУ, д.т.н., e-mail: Liliya.demidova@rambler.ru;*

*Тишкин Р.В., начальник отдела филиала ФГУП «ГНПРКЦ «ЦСКБ-Прогресс» – ОКБ «Спектр», к.т.н., e-mail: roman.tishkin@mail.ru;*

*Труханов С.В., главный специалист отдела филиала ФГУП «ГНПРКЦ «ЦСКБ-Прогресс» – ОКБ «Спектр», e-mail: serge\_tsv@mail.ru*

**Ключевые слова:** алгоритм идентификации, гиперспектральная характеристика объекта, мера подобия евклидова расстояния, нечеткая мера подобия, угловая мера подобия, нечеткая линейная регрессия, консолидация.

### Введение

В последние годы одной из важных задач, решаемых системами дистанционного зондирования Земли (ДЗЗ), является задача обработки и анализа гиперспектральной информации, которая формируется на основе полученных с борта космического аппарата гиперспектральных изображений (ГСИ) земной поверхности, представленных большим набором снимков одной и той же сцены в узких соприкасающихся спектральных диапазонах. Так, например, отечественный космический аппарат «Ресурс-П» с гиперспектральной аппаратурой (ГСА) на борту, запущенный 25 июня 2013 года, выдает наборы снимков в 130 узких соприкасающихся диапазонах видимой области спектра, которые и формируют гиперспектральное изображение (ГСИ), называемое также «гиперкубом».

Большое внимание при обработке ГСИ уделяется проблеме идентификации объектов земной поверхности по их гиперспектральным характеристикам (ГСХ). ГСХ объекта в графическом виде представляет собой отображение связи между длиной волны и значениями коэффициента спектральной яркости (КСЯ) анализируемого объекта. Идентификация объекта по его ГСХ может быть осуществлена посредством сравнения с применением тех или иных мер подобия (сходства) анализируемой ГСХ объекта с некоторыми эталонными ГСХ (эталонами), объединенными в специальные спектральные библиотеки [1–3]. При этом необходимо учитывать, что у анализируемой и эталонных ГСХ должны быть одинаковыми единицы измерения, спектральный диапазон и разрешение данных [3].

Цель работы заключается в создании и исследовании алгоритмов идентификации ГСХ объектов земной поверхности с применением различных обоснованно выбранных мер подобия, а также – в последующей консолидации частных результатов идентификации ГСХ объектов, полученных при использовании предлагаемых алгоритмов.

*Рассматривается подход к решению задачи идентификации объектов земной поверхности, основанный на анализе гиперспектральных характеристик объектов, получаемых с обработанных космических изображений, посредством применения алгоритмов идентификации на основе различных мер подобия. Представлены результаты обработки гиперспектральной информации с использованием предлагаемых алгоритмов идентификации.*

### Алгоритмы идентификации: теоретические исследования

Анализ работ отечественных и зарубежных авторов [4–8] свидетельствует о том, что в настоящее время для решения задач идентификации объектов земной поверхности по их ГСХ наиболее часто используются подходы, основанные на применении: мер подобия (сходства) [5, 9]; метода спектрального угла [4, 5, 10]; искусственных нейронных сетей [5, 12]; алгоритма  $k$ -средних [4–7, 10]; метода максимального правдоподобия (к сожалению, оказавшегося неэффективным при работе с 200-полосными ГСИ) [4, 5].

Результаты экспериментальных исследований по оценке качества идентификации объектов земной поверхности на космических снимках с применением перечисленных выше подходов, свидетельствуют о том, что ни один из них не обладает явным преимуществом перед другими, так как не обеспечивает требуемое качество идентификации объектов в преобладающем большинстве случаев.

Для решения задачи идентификации объектов земной поверхности по их ГСХ предлагается одновременно использовать четыре алгоритма идентификации ГСХ, основанные на мере подобия (сходства) евклидова расстояния, угловой мере подобия и двух нечетких мерах подобия, с целью последующей консолидации частных результатов идентификации.

Выбор меры подобия евклидова расстояния обусловлен высокой эффективностью применения данной меры при решении широкого спектра задач идентификации (классификации), в том числе и задач идентификации при обработке космических изображений.

Использование угловой меры подобия, предполагающей реализацию метода спектрального угла (Spectral Angle Mapper – SAM), может быть обосновано тем, что

данная мера обеспечивает получение хороших результатов идентификации ГСХ объектов, имеющих схожие значения яркости точек изображения во всех спектральных диапазонах даже в том случае, когда на исходных изображениях наблюдаются эффекты засветки, поскольку метод спектрального угла не учитывает значения яркости точек изображения.

Целесообразность использования нечетких мер подобия [6, 11, 13] может быть обоснована тем, что некоторая анализируемая ГСХ объекта может одновременно принадлежать к разным классам (категориям) объектов, в то время как алгоритмы «жесткой» идентификации [5], которыми являются алгоритм на основе меры подобия евклидова расстояния и алгоритм на основе угловой меры подобия, однозначно отнесут анализируемую ГСХ объекта к тем или иным однозначно определенным классам, что может быть не всегда верно. Применение алгоритмов «мягкой» идентификации, основанных на применении нечетких мер подобия, позволит более гибко решать задачу идентификации объекта по его ГСХ.

Программная реализация предлагаемых алгоритмов идентификации объекта по его ГСХ (алгоритмов сравнения ГСХ объекта с эталонами на основе меры подобия евклидова расстояния, угловой меры подобия и двух нечетких мер подобия) была выполнена при разработке системы интеллектуальной обработки данных гиперспектральной съемки (СИОДГС), структурно состоящей из базы данных хранения и обработки данных (БД) и программы интеллектуальной обработки данных гиперспектральной съемки [14, 15].

В БД СИОДГС осуществляется хранение и обработка анализируемых ГСХ; в частности, с помощью хранимых процедур проводятся операции нормирования анализируемой ГСХ к эталонным значениям, интерполяция входных значений, расчет меры подобия евклидова расстояния, угловой меры подобия, двух нечетких мер подобия, а также – обработка и хранение результатов. В программе интеллектуальной обработки данных гиперспектральной съемки СИОДГС реализованы четыре алгоритма идентификации ГСХ объекта: алгоритм идентификации ГСХ на основе меры подобия евклидова расстояния, алгоритм идентификации ГСХ на основе угловой меры подобия, два алгоритма идентификации ГСХ на основе нечетких мер подобия, а также алгоритм консолидации частных результатов идентификации ГСХ объекта.

#### **Алгоритм идентификации ГСХ на основе меры подобия евклидова расстояния**

Алгоритм идентификации ГСХ на основе меры подобия евклидова расстояния реализует идентификацию ГСХ с использованием известной евклидовой метрики для вычисления расстояния между двумя точками в  $J$ -мерном пространстве [9]:

$$E = \sqrt{\sum_{j=1}^J (y_j^A - y_j^S)^2}, \quad (1)$$

где  $y_j^A$  – значение коэффициента спектральной яркости (КСЯ) анализируемой (Analyzed) ГСХ для  $j$ -го канала гиперспектрометра;  $y_j^S$  – значение КСЯ эталонной

(Standard) ГСХ для  $j$ -го канала гиперспектрометра;  $j = \overline{1, J}$ ;  $J$  – количество каналов гиперспектрометра, равное количеству точек в анализируемой (эталонной) ГСХ (например,  $J = 96$ ).

Следует отметить, что длина волны  $\lambda_j$  для  $j$ -го ( $j = \overline{1, J}$ ) канала гиперспектрометра считается известной величиной, которой в соответствие ставятся значения КСЯ анализируемой и эталонной ГСХ.

Алгоритм идентификация ГСХ на основе меры подобия евклидова расстояния предполагает:

- расчет мер подобия евклидова расстояния (1) для анализируемой ГСХ и эталонных ГСХ, хранимых в БД;
- упорядочение по возрастанию вычисленных значений мер подобия евклидова расстояния;
- выбор в качестве искомой той эталонной ГСХ, для которой значение меры подобия евклидова расстояния (1) является минимальным.

#### **Алгоритм идентификации ГСХ на основе угловой меры подобия**

Алгоритм идентификации ГСХ на основе угловой меры подобия предполагает использование метода спектрального угла (Spectral Angle Mapper – SAM), реализующего оценку подобия исследуемой и эталонной ГСХ, рассматриваемых как векторы, размерность которых равна количеству каналов гиперспектрометра, посредством вычисления угла между ними [10]:

$$\alpha = \arccos \frac{G \cdot G'}{|G| \cdot |G'|} = \arccos \frac{\sum_{j=1}^J g_j \cdot g'_j}{\sqrt{\sum_{j=1}^J g_j^2} \cdot \sqrt{\sum_{j=1}^J (g'_j)^2}}, \quad (2)$$

где  $G$  и  $G'$  – спектры анализируемой и эталонной ГСХ соответственно;  $g_j$  и  $g'_j$  – значения КСЯ анализируемой и эталонной ГСХ соответственно для значения длины волны  $\lambda_j$  ( $j = \overline{1, J}$ );  $J$  – количество каналов гиперспектрометра.

Алгоритм идентификация ГСХ на основе угловой меры подобия предполагает:

- расчет угловых мер подобия (2) для анализируемой ГСХ и эталонных ГСХ, хранимых в БД;
- упорядочение по возрастанию вычисленных значений угловых мер подобия;
- выбор в качестве искомой той эталонной ГСХ, для которой значение угловой меры подобия (2) является минимальным.

Довольно часто из-за погрешностей ГСА, приводящих к неточности информации об анализируемой ГСХ, а также из-за почти полного отсутствия «чистых» анализируемых ГСХ и наличия в большинстве случаев анализируемых ГСХ, представляющих собой смесь из нескольких базовых классов [5], целесообразно рассматривать несколько эталонных ГСХ, имеющих наименьшие значения меры подобия евклидова расстояния (1) и угловой меры подобия (2), в качестве потенциально искомых.

Несмотря на использование при идентификации большого количества эталонных ГСХ, результаты идентификации с применением меры подобия евклидова

расстояния или угловой меры подобия может быть неудовлетворительным. В связи с этим целесообразно осуществлять подтверждение результата идентификации, которое, в частности, может быть получено посредством применения других алгоритмов идентификации ГСХ с последующей консолидацией частных результатов идентификации. Как показывают экспериментальные исследования, хороший консолидирующий результат идентификации обеспечивает совместное использование алгоритма идентификации ГСХ на основе меры подобия евклидова расстояния, алгоритма идентификации ГСХ на основе угловой меры подобия и двух алгоритмов идентификации на основе нечетких мер подобия.

### Алгоритмы идентификации ГСХ на основе нечетких мер подобия

Идентификация ГСХ объекта может быть выполнена на основе уравнения классической линейной регрессии (КЛР) посредством решения задач подбора параметров  $k$  и  $b$  уравнения (например, в соответствии с методом наименьших квадратов [16]) для анализируемой и эталонных ГСХ:

$$y = k \cdot x + b \quad (3)$$

с последующим вычислением среднеквадратичных отклонений (СКО) невязок анализируемой и эталонных ГСХ, которые могут использоваться как характеристики уникальности при идентификации ГСХ.

Если для идентификации ГСХ объекта использовать только одну характеристику уникальности – такую, как СКО, то в ряде случаев ГСХ объектов, имеющие приблизительно равные значения СКО, могут быть отнесены к одному классу, несмотря на то, что форма кривых ГСХ объектов будет различна [17]. В связи с этим для идентификации ГСХ объекта предлагается использовать дополнительные характеристики уникальности, которые могут быть получены на основе уравнения нечеткой линейной регрессии (НЛР) и использованы для расчета нечетких мер подобия.

В [18] показано, что наиболее целесообразно использовать уравнение НЛР с асимметричными нечеткими параметрами, так как оно (в отличие от уравнения нечеткой линейной регрессии с симметричными нечеткими параметрами) обеспечивает вычисление значения характеристики уникальности – СКО, которое равно или близко значению, получаемому с помощью уравнения КЛР:

$$Y(x) = A_1 \cdot x + A_0, \quad (4)$$

где  $A_1 = (a_1, c_1, d_1)$  и  $A_0 = (a_0, c_0, d_0)$  – ТНЧ, соответствующие параметрам  $k$  и  $b$  уравнения КЛР (3), представляемые с помощью треугольных функций принадлежности (подробно рассмотренных в [17]), и являющиеся асимметричными нечеткими параметрами уравнения НЛР (4).

При разработке алгоритмов идентификации ГСХ объекта на основе нечетких мер могут использоваться различные нечеткие меры подобия [11], при этом сами алгоритмы идентификации будут иметь одинаковые этапы реализации. Укрупненная схема алгоритма идентификации ГСХ объекта на основе нечеткой меры подобия приведена на рис. 1.

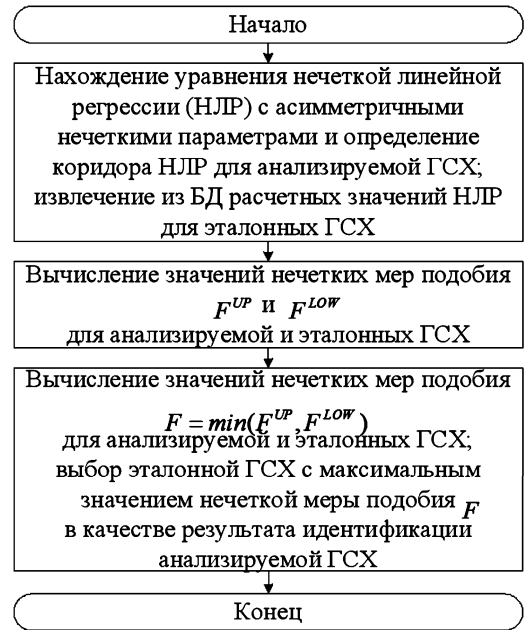


Рис. 1. Укрупненная схема алгоритма идентификации ГСХ объекта на основе нечеткой меры подобия

На первом этапе алгоритма определяются параметры (а именно ТНЧ) уравнения НЛР анализируемой ГСХ, для чего решается задача квадратичного программирования (ЗКП), формулируемая следующим образом [17, 18]:

$$F_{\text{ЗКП}} = k_1 \cdot \sum_{j=1}^J (y_j - \sum_{i=0}^n a_i \cdot x_{ji})^2 + k_2 \cdot (1 - \alpha) \times \quad (5)$$

$$\times \sum_{j=1}^J \sum_{i=0}^n (c_i + d_i) \cdot x_{ji} + \xi \cdot \sum_{i=0}^n (c_i^2 + d_i^2) \rightarrow \min_{a, c, d}$$

при ограничениях:

$$\sum_{i=0}^n a_i \cdot x_{ji} + (1 - \alpha) \cdot \sum_{i=0}^n d_i \cdot x_{ji} \geq y_j; \quad (6)$$

$$\sum_{i=0}^n a_i \cdot x_{ji} - (1 - \alpha) \cdot \sum_{i=0}^n c_i \cdot x_{ji} \leq y_j; \quad (7)$$

$$c_i \geq 0; d_i \geq 0 \quad (j = \overline{1, J}; i = \overline{0, n}; n = 1), \quad (8)$$

где  $j$  – порядковый номер канала гиперспектрометра ( $j = \overline{1, J}$ );  $J$  – количество каналов гиперспектрометра;

$i$  – порядковый номер ТНЧ ( $i = \overline{0, 1}$  при  $n = 1$ );  $x_{j0} = 1$

( $j = \overline{1, J}$ ), так как соответствуют свободному члену – параметру  $b$  – уравнения КЛР (2);  $x_{j1}$  – значение длины волны для  $j$ -го канала гиперспектрометра;  $y_j$  – значение КСЯ ГСХ для длины волны  $x_{j1}$  ( $x_{j1} = \lambda_j$ );  $\alpha$  – значение уровня ТНЧ, характеризующее ширину коридора НЛР ( $\alpha \in [0, 1]$ );  $k_1, k_2$  – весовые коэффициенты, характеризующие вклад первого и второго слагаемых в целевую функцию (4);  $\xi$  – малое положительное число такое, что  $k_1, k_2 \gg \xi$  (третье слагаемое введено в целевую функцию для того, чтобы она имела квадратичный вид и чтобы при поиске значений ТНЧ можно было сформулировать ЗКП [18]).

При решении ЗКП (5)–(8) предполагается, что  $k_1 = k_2 = 1$  (в [17] показано, что выбор значений пара-

метров  $k_1, k_2$  не оказывает сильного влияния на решение задачи);  $\xi = 0,001$ .

Пусть  $a_0 = z_1; c_0 = z_2; d_0 = z_3; a_1 = z_4; c_1 = z_5; d_1 = z_6; x_{j1} = \lambda_j; y_j = g_j$ . Тогда ЗКП (5)–(8) может быть записана как:

$$F_{\text{ЗКП}} = \sum_{j=1}^J (g_j - z_1 - z_4 \cdot \lambda_j)^2 + \sum_{j=1}^J (z_2 + z_5 \cdot \lambda_j + z_6 \cdot \lambda_j) + 0,001 \cdot (z_2^2 + z_3^2 + z_5^2 + z_6^2) \rightarrow \min_{z_1, \dots, z_6} \quad (9)$$

при ограничениях:

$$\begin{cases} z_1 + z_3 + z_4 \cdot \lambda_j + z_6 \cdot \lambda_j \geq g_j; \\ z_1 - z_2 + z_4 \cdot \lambda_j - z_5 \cdot \lambda_j \leq g_j; \\ z_2 \geq 0; \\ z_3 \geq 0; \\ z_5 \geq 0; \\ z_6 \geq 0. \end{cases} \quad (10)$$

Для решения ЗКП в виде (8) при ограничениях (9) может быть использован метод неопределенных множителей Лагранжа [16]. Ввиду того, что решение ЗКП методом множителей Лагранжа является весьма трудоемкой задачей, целесообразно использование существующих математических пакетов, позволяющих решать задачи минимизации с ограничениями. В частности, ЗКП может быть решена в системе инженерных и научных расчетов MATLAB с применением встроенной функции «quadprog» [19].

Уравнение НЛР (3) для анализируемой ГСХ объекта строится на основе решения ЗКП (9)–(10). При этом на основе вычисленных ТНЧ  $A_1 = (a_1, c_1, d_1)$  и  $A_0 = (a_0, c_0, d_0)$  уравнения НЛР (4) для значений длин волн  $x = \lambda_j (j = \overline{1, J})$  определяются характерные точки уравнения КЛР  $Y_{\text{НЛР}}^{\text{КЛР}}(\lambda_j)$ :

$$Y_{\text{НЛР}}^{\text{КЛР}}(\lambda_j) = a_0 + a_1 \cdot \lambda_j, \quad (11)$$

а также уравнений верхней  $Y_{\text{НЛР}}^{\text{UP}}(\lambda_j)$  и нижней  $Y_{\text{НЛР}}^{\text{LOW}}(\lambda_j)$  границ коридора НЛР анализируемой характеристики:

$$Y_{\text{НЛР}}^{\text{UP}}(\lambda_j) = a_0 + d_0 + (a_1 + d_1) \cdot \lambda_j, \quad (12)$$

$$Y_{\text{НЛР}}^{\text{LOW}}(\lambda_j) = a_0 - c_0 + (a_1 - c_1) \cdot \lambda_j. \quad (13)$$

Для эталонных ГСХ уравнения НЛР (3), а также уравнения КЛР  $Y_{\text{НЛР}}^{\text{КЛР}}(\lambda_j)$ , верхней  $Y_{\text{НЛР}}^{\text{UP}}(\lambda_j)$  и нижней  $Y_{\text{НЛР}}^{\text{LOW}}(\lambda_j)$  границ коридора НЛР, определяемые в соответствии с (11), (12) и (13), строятся аналогичным образом, при этом все расчетные значения НЛР для эталонных ГСХ хранятся в БД СИОДГС.

Из-за асимметричности ТНЧ  $A_1 = (a_1, c_1, d_1)$  и  $A_0 = (a_0, c_0, d_0)$  коридоры НЛР анализируемой и эталонных ГСХ также асимметричны (рис. 2). В связи с этим точки ГСХ (как анализируемой, так и каждой эталонной) могут быть разбиты на 2 подмножества: подмножество точек, лежащих в верхней части *UP* коридора НЛР (между линией верхней границы коридора НЛР и линией КЛР, определяемыми соответственно уравнениями (12) и (11)), и подмножество точек, лежащих в нижней части *LOW* коридора НЛР (между линией нижней границы

коридора НЛР и линией КЛР, определяемыми соответственно уравнениями (13) и (11)).

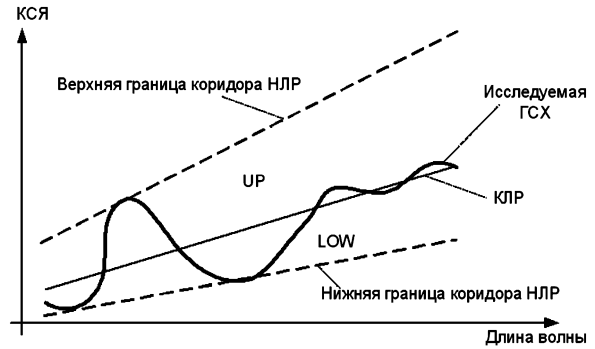


Рис. 2. Представление ГСХ в коридоре НЛР

На втором этапе алгоритма для точек анализируемой и каждой эталонной ГСХ, лежащих в верхней *UP* и нижней *LOW* частях коридоров своих НЛР, находятся значения нечетких мер подобия  $F^{UP}$  и  $F^{LOW}$ , для вычисления которых используется одна из двух нечетких мер подобия вида [11, 17]:

$$f_1 = 1 - \frac{\sum_{j=1}^J |u_A(\lambda_j, g'_j) - u_S(\lambda_j, g''_j)|}{\sum_{j=1}^J (u_A(\lambda_j, g'_j) + u_S(\lambda_j, g''_j))}, \quad (14)$$

$$f_2 = \frac{1}{J} \cdot \sum_{j=1}^J \frac{\min(u_A(\lambda_j, g'_j), u_S(\lambda_j, g''_j))}{\max(u_A(\lambda_j, g'_j), u_S(\lambda_j, g''_j))}, \quad (15)$$

где  $u_A(\lambda_j, g'_j)$  – значение функции принадлежности нечеткого множества *A* (Analyzed) анализируемой ГСХ уравнению НЛР этой ГСХ для значения КСЯ  $g'_j$ , соответствующего длине волны  $\lambda_j (j = \overline{1, J})$ ;  $u_S(\lambda_j, g''_j)$  – значение функции принадлежности нечеткого множества *S* (Standard) эталонной ГСХ уравнению НЛР этой ГСХ для значения КСЯ  $g''_j$ , соответствующего длине волны  $\lambda_j (j = \overline{1, J})$ ; *J* – количество каналов гиперспектрометра, равное количеству точек в анализируемой (эталонной) ГСХ.

Данные меры подобия были выбраны из множества известных нечетких мер подобия как показавшие наивысшее качество результатов идентификации при выполнении оценки подобия тестовых наборов данных, описываемых с помощью уравнений НЛР [17].

Значение функции принадлежности некоторой точки  $(\lambda_j, g_j)$  ГСХ к уравнению НЛР этой ГСХ может быть определено как [17, 18]:

$$u(\lambda_j, g_j) = \begin{cases} 1 - \frac{a_0 + a_1 \cdot \lambda_j - g_j}{c_0 + c_1 \cdot \lambda_j}, & \text{если} \\ a_0 + a_1 \cdot \lambda_j - c_0 - c_1 \cdot \lambda_j \leq g_j \leq a_0 + a_1 \cdot \lambda_j; \\ 1 - \frac{g_j - a_0 - a_1 \cdot \lambda_j}{d_0 + d_1 \cdot \lambda_j}, & \text{если} \\ a_0 + a_1 \cdot \lambda_j \leq g_j \leq a_0 + a_1 \cdot \lambda_j + d_0 + d_1 \cdot \lambda_j; \\ 0, & \text{иначе;} \end{cases} \quad (16)$$

где  $a_0, c_0, d_0, a_1, c_1, d_1$  – параметры ТНЧ, вычисленные при решении ЗКП (9) – (10).

На завершающем **третьем этапе** алгоритма для анализируемой и каждой эталонной ГСХ сначала проводится вычисление результирующей нечеткой меры подобия, определяемой как минимум из нечетких мер подобия  $F^{UP}$  и  $F^{LOW}$ :

$$F = \min(F^{UP}, F^{LOW}), \quad (17)$$

а затем все эталонные ГСХ упорядочиваются по убыванию вычисленных таким образом значений нечетких мер подобия, при этом в качестве искомой выбирается та эталонная ГСХ, для которой значение нечеткой меры подобия (17) является максимальным.

#### **Алгоритм консолидации частных результатов идентификации ГСХ объекта**

Как уже было отмечено, для повышения качества идентификации анализируемой ГСХ объекта целесообразно выполнение консолидации частных результатов идентификации (тем или иным образом) [20].

При применении четырех алгоритмов идентификации ГСХ к анализируемой ГСХ и некоторой эталонной ГСХ вычисляются четыре значения: значение меры сходства евклидова расстояния  $E$  (1); значение угловой меры подобия  $\alpha$  (2); два значения нечетких мер подобия  $F_1$  и  $F_2$ , вычисленных соответственно с применением формул (14), (15) и (17).

При работе с БД СИОДГС, содержащей информацию об  $K$  эталонных ГСХ, для анализируемой ГСХ объекта будут получены четыре идентифицирующих набора, каждый из которых является результатом применения одного из четырех алгоритмов идентификации ГСХ и содержит  $K$  значений (в соответствии с количеством эталонных ГСХ в БД). Каждый такой идентифицирующий набор может быть упорядочен по убыванию (возрастанию) значений используемой меры подобия, в результате чего эталонным ГСХ из БД будут присвоены некоторые рейтинговые оценки (порядковые номера в рейтинге): чем выше значения меры подобия евклидова расстояния и угловой меры сходства и чем ниже значения нечетких мер подобия, тем меньше номер в рейтинге.

Пусть  $R_E^k$  – рейтинговая оценка  $k$ -й эталонной ГСХ при использовании алгоритма идентификации ГСХ на основе меры подобия евклидова расстояния  $E$ ;  $R_\alpha^k$  – рейтинговая оценка  $k$ -й эталонной ГСХ при использовании алгоритма идентификации ГСХ на основе угловой меры подобия  $\alpha$ ;  $R_{F_1}^k$  – рейтинговая оценка  $k$ -й эталонной ГСХ при использовании алгоритма идентификации ГСХ на основе нечеткой меры подобия  $F_1$ ;  $R_{F_2}^k$  – рейтинговая оценка  $k$ -й эталонной ГСХ при использовании алгоритма идентификации ГСХ на основе нечеткой меры подобия  $F_2$  ( $k = \overline{1, K}$ ).

Алгоритм консолидации частных результатов идентификации ГСХ объекта предполагает:

консолидацию частных результатов идентификации, полученных с применением меры сходства евклидова расстояния  $E$ , угловой меры подобия  $\alpha$ , а также двух

нечетких мер подобия  $F_1$  и  $F_2$  посредством применения формулы:

$$\bar{R}^k = \frac{(R_E^k + R_\alpha^k + R_{F_1}^k + R_{F_2}^k)}{4}; \quad (18)$$

упорядочение эталонных ГСХ из БД по возрастанию усредненных значений рейтинговых оценок  $\bar{R}^k$  ( $k = \overline{1, K}$ ).

Пусть консолидирующие рейтинговые оценки  $\hat{R}^k$  определены как номера в рейтинге для усредненных оценок  $\bar{R}^k$  ( $k = \overline{1, K}$ ). Очевидно, что для дальнейшего анализа целесообразно использовать незначительную часть лучших эталонных ГСХ из БД (от 5 до 10 эталонных ГСХ), имеющих наименьшие значения (порядковые номера) консолидирующих рейтинговых оценок  $\hat{R}^k$ .

После визуального сравнительного анализа результатов идентификации анализируемой ГСХ объекта с помощью некоторого количества лучших эталонных ГСХ из БД, полученных вышеуказанным образом и представленных как в текстовом, так и в графическом виде, оператор программы СИОДГС может принять окончательное решение о соответствии анализируемой ГСХ объекта некоторой эталонной ГСХ из БД.

#### **Экспериментальные исследования**

Предлагаемая СИОДГС была использована при решении задачи идентификации ГСХ различных объектов, в частности, была проанализирована ГСХ «Сосна, крона, Московская обл., 1 августа, 1970 г.», взятая с сайта GIS-LAB [3]. В БД на момент выполнения данного анализа находилось 281 эталонная ГСХ, включая анализируемую ГСХ.

При решении задачи идентификации анализируемой ГСХ с применением четырех рассмотренных выше алгоритмов идентификации были использованы все эталонные ГСХ из БД СИОДГС. В табл. 1 представлены пять лучших результатов идентификации анализируемой ГСХ, при этом указаны порядковые номера  $\hat{R}^k$  лучших эталонных ГСХ, определенные посредством упорядочения по возрастанию консолидирующих рейтинговых оценок  $\bar{R}^k$  ( $k = \overline{1, K}$ ) всех эталонных ГСХ (при идентификации анализируемой ГСХ четырьмя алгоритмами идентификации), наименования лучших эталонных ГСХ с указанием их класса и типа, а также порядковые номера приведенных в таблице эталонных ГСХ, определенные при индивидуальном применении алгоритмов идентификации ГСХ к ГСХ «Сосна, крона, Московская обл., 1 августа, 1970 г.» на основе меры подобия евклидова расстояния, угловой меры подобия и нечетких мер подобия ( $R_E, R_\alpha, R_{F_1}, R_{F_2}$ ). Следует отметить, что во всех представленных ниже таблицах, в том числе и в таблице 1, жирным шрифтом выделены порядковые номера трех лучших ГСХ, определенные как при индивидуальном применении алгоритмов идентификации ГСХ, так и в результате консолидации частных результатов идентификации ГСХ.

На рис. 3 и 4 приведены фрагменты главного окна программы, содержащие графическую интерпретацию результатов идентификации ГСХ. При этом показаны: анализируемая ГСХ, линия КЛР, границы коридора НЛР

и эталонная ГСХ (выбранная из таблицы результатов идентификации ГСХ). Точки анализируемой ГСХ изображены темными квадратами, точки эталонной ГСХ – светлыми окружностями.

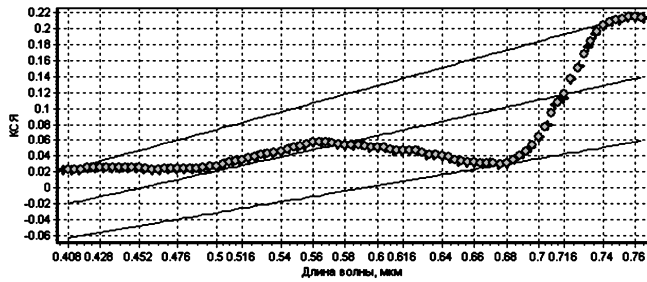


Рис. 3. Фрагмент главного окна программы с результатом идентификации, представленным в графическом виде (для эталона с общим порядковым номером, равным 1)

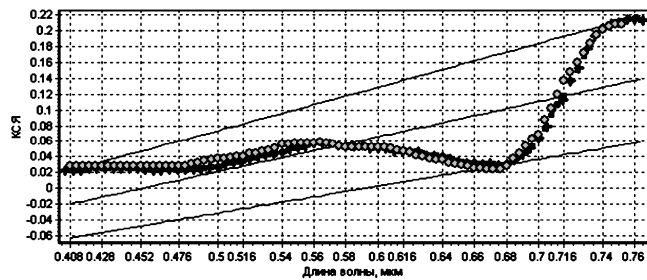


Рис. 4. Фрагмент главного окна программы с результатом идентификации, представленным в графическом виде (для эталона с общим порядковым номером, равным 2)

На рис. 3 представлены анализируемая ГСХ и эталонная ГСХ с порядковым номером 1 (табл. 1), соответствующей объекту «Сосна, крона, Московская обл., 1 августа, 1970 г.». При этом все точки анализируемой ГСХ полностью совпадают с точками эталонной ГСХ. На

рис. 4 представлены анализируемая ГСХ и эталонная ГСХ с порядковым номером 2 (табл. 1), соответствующей объекту «Сосна, крона, Московская обл., 6 августа, 1970 г.». При этом можно увидеть наличие некоторых расхождений между анализируемой и эталонной ГСХ.

Из табл. 1 видно, что все эталонные ГСХ лучших результатов идентификации принадлежат к тому же классу («Деревья») и к тому же типу («Растительность»), что и анализируемая ГСХ «Сосна, крона, Московская обл., 1 августа, 1970 г.» (хотя все эталонные ГСХ, участвовавшие в идентификации, разделены на 4 типа и 21 класс).

Следует отметить, что при использовании только одного алгоритма идентификации ГСХ, результат идентификации может быть менее точным и даже неверным. Так, например, из табл. 2, содержащей 5 лучших результатов идентификации анализируемой ГСХ с помощью только одного алгоритма идентификации ГСХ на основе угловой меры подобия  $R_\alpha$ , видно, что, если бы БД не содержала эталонную ГСХ «Сосна, крона, Московская обл., 1 августа, 1970 г.», то результат идентификации – выбранная в качестве искомой эталонная ГСХ – мог бы иметь неправильно определенные класс и тип анализируемой ГСХ. В то же время консолидация частных результатов идентификации ГСХ, полученных с применением четырех алгоритмов идентификации, позволяет

установить более точные рейтинговые оценки  $\bar{R}^k$  эталонных ГСХ из БД и избежать принятия ложного решения о результатах идентификации анализируемой ГСХ (при отсутствии одноименного эталона ГСХ в БД) как ГСХ «Чернозем типичный, гор. А1 (0 - 50 см), 1975 г.» (табл. 2, вторая строка). При этом анализируемая ГСХ была бы идентифицирована как ГСХ «Сосна, крона, Московская обл., 6 августа, 1970 г.» (табл. 1, вторая строка) с правильно определенными классом и типом.

Таблица 1. Лучшие результаты идентификации ГСХ «Сосна, крона, Московская обл., 1 августа, 1970 г.» (консолидация)

Порядк. номер $\hat{R}^k$	Наименование объекта с эталонной ГСХ	Класс	Тип	Среднее арифм. $\bar{R}^k$	$R_E$	$R_\alpha$	$R_{F_1}$	$R_{F_2}$
1	Сосна, крона, Московская обл., 1 августа, 1970 г.	Деревья	Растительность	1	1	1	1	1
2	Сосна, крона, Московская обл., 6 августа, 1970 г.	Деревья	Растительность	6,25	4	8	10	3
3	Сосна, крона, Московская обл., 23 августа, 1970 г.	Деревья	Растительность	7	10	11	5	2
4	Береза, крона, Московская обл., 6 августа, 1970 г.	Деревья	Растительность	14	3	10	24	19
5	Береза, крона, Московская обл., 14 июля, 1970 г.	Деревья	Растительность	14,25	18	12	15	12

Таблица 2. Лучшие результаты идентификации ГСХ «Сосна, крона, Московская обл., 1 августа, 1970 г.» на основе угловой меры подобия (2)

$R_\alpha$	Наименование объекта с эталонной ГСХ	Класс	Тип	Порядк. номер $\hat{R}^k$	Среднее арифм. $\bar{R}^k$	$R_E$	$R_{F_1}$	$R_{F_2}$
1	Сосна, крона, Московская обл., 1 августа, 1970 г.	Деревья	Растительность	1	1	1	1	1
2	Чернозем типичный, гор. А1 (0 - 50 см), 1975 г.	Почвы	Грунты	98	24,5	15	219	200
3	Дерново-карбонатная почва. А1 (0 - 12 см), 1975 г.	Почвы	Грунты	153	38,25	74	280	280
4	Дерново-карбонатная почва А1 (>12 см), 1975 г.	Почвы	Грунты	203	50,75	208	281	281
5	Сосна, крона, Московская обл., 24 июня (молодые побеги), 1970 г.	Деревья	Растительность	25	6,25	136	2	7

Таблица 3. Лучшие результаты идентификации ГСХ «Асфальтированное шоссе сухое, 17 июля, 1960 г.» на основе угловой меры подобия (2)

$R_{\alpha}$	Наименование объекта с эталонной ГСХ	Класс	Тип	Порядк. номер $\hat{R}^k$	Среднее арифм. $\bar{R}^k$	$R_E$	$R_{F_1}$	$R_{F_2}$
1	Асфальтированное шоссе сухое, 17 июля, 1960 г.	Сооружения	Антропогенные объекты	1	1	1	1	1
2	Чернозем обыкновенный, почва суббореального пояса	почвы	Грунты	52	87,25	108	85	104
3	Шоссе с асфальтовым покрытием, сухое, 17 июля 1961 г.	сооружения	Антропогенные объекты	2	4,25	2	7	6
4	Дерново-карбонатная почва. А1 (0 - 12 см).	почвы	Грунты	141	173,75	124	231	199
5	Каолин	горные породы	Грунты	61	95	278	20	21

Таблица 4. Лучшие результаты идентификации ГСХ «Асфальтированное шоссе сухое, 17 июля, 1960 г.» на основе нечеткой меры подобия (14)

$R_{F_1}$	Наименование объекта с эталонной ГСХ	Класс	Тип	Порядк. номер $\hat{R}^k$	Среднее арифм. $\bar{R}^k$	$R_{\alpha}$	$R_E$	$R_{F_2}$
1	Асфальтированное шоссе сухое, 17 июля, 1960 г.	Сооружения	Антропогенные объекты	1	1	1	1	1
2	Средние горные породы: трахит	Горные породы	Грунты	46	79,25	18	251	2
3	Береза, крона, Московская обл., 1 октября 1970 г.	Деревья	Растительность	88	152,75	256	264	3
4	Такыр типичный, распаханная почва субтропического пояса	Почвы	Грунты	58	91,5	65	236	7
5	Грубообломочные горные породы: желтовато-серый мелкогалечный конгломерат	Горные породы	Грунты	4	28,5	84	22	4

Таблица 5. Лучшие результаты идентификации ГСХ «Асфальтированное шоссе сухое, 17 июля, 1960 г.» на основе меры подобия евклидова расстояния (1)

$R_E$	Наименование объекта с эталонной ГСХ	Класс	Тип	Порядк. номер $\hat{R}^k$	Среднее арифм. $\bar{R}^k$	$R_{\alpha}$	$R_{F_1}$	$R_{F_2}$
1	Асфальтированное шоссе сухое, 17 июля, 1960 г.	Сооружения	Антропогенные объекты	1	1	1	1	1
2	Шоссе с асфальтовым покрытием, сухое, 17 июля 1961 г.	Сооружения	Антропогенные объекты	2	4,5	3	7	6
3	Почва: солонец мелкий	Почвы	Грунты	15	44	48	61	52
4	Песчаные подзолистые почвы. Содержание влаги в почве - 0-4%	Почвы	Грунты	43	76,25	21	10 8	133
5	Шоссе, мощенное камнем, сухое, 17 июля 1960 г.	Сооружения	Антропогенные объекты	5	30	16	46	53

Алгоритмы идентификации ГСХ на основе угловой меры подобия и нечетких мер подобия также не всегда дают приемлемые по точности результаты идентификации ГСХ [в частности, при идентификации ГСХ сухого асфальтированного шоссе алгоритмы идентификации ГСХ на основе угловой меры подобия и нечетких мер подобия дают менее точные результаты идентификации, чем алгоритм идентификации ГСХ на основе меры сходства евклидова расстояния (табл. 3–5)].

Так, например, из табл. 3 и 4, содержащих 5 лучших результатов идентификации ГСХ «Асфальтированное шоссе сухое, 17 июля, 1960 г.» соответственно с помощью только одного алгоритма идентификации ГСХ на основе угловой меры подобия (2)  $R_{\alpha}$  и только одного алгоритма идентификации ГСХ на основе нечеткой меры (14)  $R_{F_1}$  видно, что, если бы БД не содержала эталонную ГСХ «Асфальтированное шоссе сухое, 17 июля, 1960 г.», то результаты идентификации – выбранные в качестве искомой эталонные ГСХ «Чернозем обыкновенный, почва суббореального пояса» (табл. 3, строка 2)

и ГСХ «Средние горные породы: трахит» (табл. 4, строка 2) – могли бы иметь неправильно определенные класс и тип анализируемой ГСХ.

В тоже время из табл. 5, содержащей 5 лучших результатов идентификации ГСХ «Асфальтированное шоссе сухое, 17 июля, 1960 г.» с помощью только одного алгоритма идентификации ГСХ на основе меры сходства евклидова расстояния (1)  $R_E$ , видно, что, даже если бы БД не содержала эталонную ГСХ «Асфальтированное шоссе сухое, 17 июля, 1960 г.», то результат идентификации – выбранная в качестве искомой эталонная ГСХ «Шоссе с асфальтовым покрытием, сухое, 17 июля 1961 г.» (табл. 5, строка 2) – имел бы, по крайней мере, правильно определенные класс и тип анализируемой ГСХ.

При этом консолидация частных результатов идентификации ГСХ «Асфальтированное шоссе сухое, 17 июля, 1960 г.», полученных с применением четырёх предлагаемых алгоритмов идентификации, позволяет принять правильное решение по идентификации анализируемой ГСХ.



Приведенные примеры свидетельствуют о целесообразности совместного применения нескольких алгоритмов идентификации ГСХ с целью последующей консолидации частных результатов идентификации.

### Заключение

Результаты экспериментальных исследований подтверждают целесообразность дальнейшего развития предлагаемого подхода к решению задачи идентификации ГСХ объектов, основанного на консолидации частных результатов идентификации ГСХ объектов, полученных с применением различных обоснованно выбранных алгоритмов идентификации ГСХ объектов, с целью повышения надежности классификационного решения. Использование такого подхода позволит решить задачу идентификации объектов земной поверхности по выделенным с обработанных космических снимков КА «Ресурс-П» гиперспектральным характеристикам с последующим накоплением эталонных характеристик в базе данных, что, в свою очередь, обеспечит создание актуальной отечественной спектральной библиотеки эталонов, которую можно будет применять при мониторинге состояния сельскохозяйственных угодий, лесных массивов, водных ресурсов, экологического состояния почв и т.п.

### Литература

1. USGS Spectroscopy Lab. <http://speclab.cr.usgs.gov>.
2. Jet Propulsion Laboratory. ASTER Spectral Library. NASA. <http://speclab.jpl.nasa.gov>.
3. Спектральные библиотеки – источник данных по спектрам. GIS-LAB. <http://www.gis-lab.info>.
4. Программный комплекс ENVI. Уч. пособие. Компания «Совзонд». – 2009. <http://www.sovzond.ru>.
5. Шовенгердт Р.А. Дистанционное зондирование. Модели и методы обработки изображений. – М.: Техносфера, 2010. – 560 с.
6. Чандра А.М., Кош С.К. Дистанционное зондирование и географические информационные системы. – М.: Техносфера, 2008. – 312 с.
7. Рис У.Г. Основы дистанционного зондирования. – М.: Техносфера, 2006. – 336 с.
8. Пылькин А.Н., Тишкин Р.В., Труханов С.В. Задачи DATA MINING и их решение в современных реляционных СУБД // Вестник Рязанского государственного радиотехнического университета. – 2011. – №4 (выпуск 38). – С. 60-65.
9. Чубукова И.А. Data Mining. Основы информационных технологий. Специальные курсы. Издательство «Бинном». Лаборатория знаний, 2006. – 384 с.
10. Yang C., Everitt J H., Bradford J.M. Yield estimation from hyperspectral imagery using spectral angle mapper (SAM). – American Society of Agricultural and Biological Engineers. Vol. 51(2): 729-737.
11. Van der Weken D., Nachtegaele M., Kerre E.E. An overview of similarity measures for images // Proceedings of

ICASSP 2002 (IEEE Int. Conf. Acoustics, Speech and Signal Processing). – Orlando, USA. – 2002. – pp. 3317-3320.

12. Пылькин А.Н., Тишкин Р.В. Методы и алгоритмы сегментации изображений. – М.: Горячая линия – Телеком, 2010. – 92 с.

13. Мятлов Г.Н., Тишкин Р.В., Ушенкин В.А., Юдаков А.А. Применение нечетких мер подобия в задаче совмещения изображений поверхности Земли // Вестник Рязанского государственного радиотехнического университета. – 2013. – №2 (выпуск 44). – С. 18-26.

14. Труханов С.В., Юдаков А.А. Создание структуры базы данных системы интеллектуальной обработки данных гиперспектральной съемки. Свидетельство о государственной регистрации программы для ЭВМ в Федеральной службе по интеллектуальной собственности №2013611036 от 09.01.2013.

15. Труханов С.В., Юдаков А.А. Программа интеллектуальной обработки данных гиперспектральной съемки. Свидетельство о государственной регистрации программы для ЭВМ в Федеральной службе по интеллектуальной собственности №2013610619 от 09.01.2013.

16. Кремер Н.Ш., Путко Б.А., Тришин И.М., Фридман М.Н. Высшая математика для экономистов: Учебник для вузов, М.: ЮНИТИ, 2002. – 471 с.

17. Демидова Л.А., Мятлов Г.Н. Подход к оценке уникальности кусочно-линейных объектов с использованием нечеткой линейной регрессии // Системы управления и информационные технологии. 2013. Т. 51. № 1. С. 85-89.

18. Haekwan Lee, Hideo Tanaka Fuzzy approximations with non-symmetric fuzzy parameters in fuzzy regression analysis. Osaka prefecture University. Journal of the Operations Research Society of Japan. Vol.42, №. 1, March 1999.

19. Ануфриев Е.И., Смирнов А.Б., Смирнова Е.Н. MATLAB 7. – СПб.: БХВ-Петербург, 2005. – 1104 с.

20. Демидова Л.А., Тишкин Р.В., Юдаков А.А. Разработка ансамбля алгоритмов кластеризации на основе матриц подобия меток кластеров и алгоритма спектральной факторизации // Вестник Рязанского государственного радиотехнического университета. 2013. № 4-1 (46). С. 9-17.

### THE OBJECTS HYPERSPECTRAL FEATURE IDENTIFICATION ALGORITHMS IN THE EARTH REMOTE SENSING TASKS

*Demidova L.A., Tishkin R.V., Trukhanov S.V.*

The approach to resolve the task of Earth objects' identification founded on hyperspectral objects' feature analyze, which are got from processed space images, by means of identification algorithms, based on the different similarity measures, is considered. The results of hyperspectral information processing by means of the offered identification algorithms, are represented.

## ПОВЫШЕНИЕ ПРОСТРАНСТВЕННОГО РАЗРЕШЕНИЯ МАТЕРИАЛОВ ГИПЕРСПЕКТРАЛЬНОЙ СЪЕМКИ ЗЕМЛИ НА ОСНОВЕ ИХ КОМПЛЕКСИРОВАНИЯ С ВЫСОКОДЕТАЛЬНЫМИ СНИМКАМИ

*Еремеев В.В., директор НИИ «Фотон» Рязанского государственного радиотехнического университета (РГРТУ), д.т.н., e-mail: foton@rsreu.ru;*  
*Макаренков А.А., аспирант НИИ «Фотон» РГРТУ;*  
*Москвитин А.Э., к.т.н., в.н.с. НИИ «Фотон» РГРТУ.*

**Ключевые слова:** гиперспектральные изображения, комплексирование, спектральное разделение, многозональные изображения, повышение пространственного разрешения.

### Введение

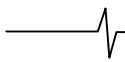
Спутниковые гиперспектральные снимки по сравнению с многозональными обладают существенно меньшим пространственным разрешением. Это накладывает ряд ограничений на области их применения. Часто проекция одного пикселя гиперспектрального изображения (ГСИ) на земной поверхности может содержать сотни пикселей, полученных многозональным датчиком более высокого пространственного разрешения [1, 2]. Так что гиперспектрометр регистрирует некоторую усредненную спектральную характеристику сравнительно большого участка Земли. Полученная в результате спектральная характеристика (СХ) описывает усредненные свойства всех СХ объектов, входящих в этот участок (так называемых спектральных составляющих), т.е. образуется «смешанная» СХ. При этом уменьшение проекции пикселя не приводит к отсутствию эффекта «смешивания» спектральных характеристик, т.к. часто объекты характеризуются смесью на уровне частиц (например, почва с минеральными примесями). В тоже время при тематической обработке гиперспектральных изображений требуется знание спектра отдельных более малых объектов, а не смеси их спектральных характеристик.

В ряде работ [3-5] рассматривается задача спектрального разделения, т.е. получения спектров отдельных объектов, составляющих пиксель гиперспектрального изображения, на основе статистической обработки самих данных. Так, в работе [3] предлагается применение метода главных компонент для определения состава и процентных долей смешанных пикселей ГСИ. В работе [5] предлагается подход, основанный на использовании линейной регрессии для поиска составляющих смешанной СХ при известном наборе эталонных («чистых») спектральных характеристик. Эти подходы для решения задачи спектрального разделения, т.е. получения спектров отдельных объектов, составляющих пиксель гиперспектрального изображения, основаны на статистической обработке самих гиперспектральных данных. Недостатком этих подходов является необходимость применения библиотек спектральных характеристик, что требует высокоточной взаим-

*Рассматриваются вопросы повышения пространственного разрешения гиперспектральных данных. Предлагается новый подход к решению этой задачи с использованием алгоритма спектрального разделения пикселей гиперспектрального изображения. Спектральное разделение предлагается выполнять с привлечением материалов синхронной многозональной высокодетальной съемки. Представлены результаты обработки натурной гиперспектральной информации.*

ной калибровки видеоданных и СХ из библиотек. Другой недостаток обусловлен сложностью точного определения процентных долей примесей в пикселе. В тоже время этот процесс может быть выполнен на более качественном уровне с использованием данных синхронной гиперспектральной и высокодетальной съемки [6-9]. В этом случае в качестве опорной информации при спектральном разделении ГСИ привлекаются материалы синхронной многозональной съемки с многократно большим пространственным разрешением по сравнению с ГСИ. То есть для анализа каждого пикселя ГСИ, которому соответствует некоторая апертура, привлекаются десятки и сотни элементов высокодетального многозонального изображения, попавших в эту апертуру. Если в упомянутых выше работах используются идеи технологии *pan-sharpening* [1, 2, 6-9], а также технологии спектрального разделения на основе привлечения спектральных библиотек [3-5], то ниже рассматривается иной подход к решению задачи повышения пространственного разрешения гиперспектральных изображений. Он основан на применении идей спектрального разделения ГСИ с привлечением высокодетальных снимков.

Пусть имеется ГСИ  $B = \{B_{mnk} \mid m = \overline{1, M}, n = \overline{1, N}, k = \overline{1, K}\}$ , где  $m, n$  – номера строк и столбцов элементов изображения,  $k$  – номер спектрального канала. Будем считать, что синхронно с ГСИ получен многозональный снимок высокого разрешения  $D = \{D_{rsp} \mid r = \overline{1, R}, s = \overline{1, S}, p = \overline{1, P}\}$ ,  $r, s$  – номера строк и столбцов элементов изображения,  $p$  – номер спектрального канала. При этом одному элементу ГСИ  $B_{mn}$  соответствует область (апертура)  $\hat{D}_{mnp}$  на снимке  $D$ :  $\hat{D}_{mnp} = \{D_{rsp} \mid r = \overline{r_1, r_2}, s = \overline{s_1, s_2}, p = \overline{1, P}\}$ ,  $r_1 = m\omega$ ,  $r_2 = (m+1)\omega$ ,  $s_1 = n\omega$ ,  $s_2 = (n+1)\omega$ , где  $\omega$  – отношение пространственных разрешений сним-



ков  $B$  и  $D$ , т.е.  $R = \omega M$ . При  $m = const$  и  $n = const$ :  $B_{mn} = (B_{mnk}, k = \overline{1, K})$  – вектор значений энергетических яркостей, образующих спектральную характеристику точки  $(m, n)$  гиперспектрального изображения. При  $r = const$  и  $s = const$ :  $D_{rs} = (D_{rsp}, p = \overline{1, P})$  – вектор значений спектральной характеристики для точки  $(r, s)$  многозонального изображения, при этом  $K \gg P$ . Каждому каналу  $k$  гиперспектрального изображения соответствует вполне определенный спектральный диапазон  $\lambda_k = [\lambda_{k1}, \lambda_{k2}]$ , где  $\lambda_{k1}$  и  $\lambda_{k2}$  – левая и правая граница спектрального диапазона с центром  $\bar{\lambda}_k$ . Аналогично определим для многозонального изображения:  $\lambda_p = [\lambda_{p1}, \lambda_{p2}]$  и центр  $\bar{\lambda}_p$ . Введем дискретную функцию  $h(\lambda)$ , ставящую в соответствие произвольной длине волны  $\lambda$  канал гиперспектрального изображения  $k = h(\lambda)$  такой, что  $\lambda \in [\lambda_{k1}, \lambda_{k2}]$ . Определим функцию сравнения спектральных характеристик многозонального и гиперспектрального изображений как  $\Phi(B_{mn}, D_{rs})$ . При этом  $\Phi(B_{m1n1}, D_{r1s1}) > \Phi(B_{m2n2}, D_{r2s2})$ , если точки  $(m_1, n_1)$  и  $(r_1, s_1)$  соответствуют однотипному сюжету подстилающей поверхности, а точки  $(m_2, n_2)$  и  $(r_2, s_2)$  – различающемуся. Если  $\Phi(B_{m1n1}, D_{r1s1}) \geq \tilde{\Phi}$ , то будем считать, что точки  $(m_1, n_1)$  и  $(r_1, s_1)$  принадлежат одному и тому же объекту. В данном случае под  $\tilde{\Phi}$  понимается минимальное значение функции  $\Phi(B_{m1n1}, D_{r1s1})$ .

**Субпиксельное спектральное разделение ГСИ**

Для каждой точки гиперспектрального изображения  $B$  определяются спектральные характеристики входящих в ее апертуру объектов и их процентные доли по площади. В результате формируется новое ГСИ  $B'$  с пространственным разрешением исходного гиперспектрального снимка  $B$ . При этом каждой точке  $B'$  приписывается несколько спектральных характеристик (по числу объектов, входящих в пиксель изображения  $B$ ). Каждая спектральная характеристика сопровождается весовым коэффициентом, характеризующим относительную площадь объекта внутри точки.

Конкретно алгоритм реализуется в виде последовательного выполнения следующих шагов.

**Шаг 1.** Формируется новое ГСИ  $\bar{B} = \{\bar{B}_{mnp} \mid m = \overline{1, M}, n = \overline{1, N}, p = \overline{1, P}\}$ , путем приведения  $B$  к спектральным диапазонам высокодетального изображения  $D$ . Эта процедура выполняется путем интегрирования значений спектральной характеристики ГСИ  $B$  в диапазонах длин волн, соответствующих многозональному изображению  $D$ :

$$\bar{B}_{mnp} = \sum_{\lambda=\lambda_{p1}}^{\lambda_{p2}} \delta_p(\lambda) B_{mnh(\lambda)},$$

где  $\delta_p(\lambda)$  – передаточная характеристика канала  $p$  изображения  $D$ .

В результате формируется новое многозональное изображение  $\bar{B}$  с тем же числом каналов  $P$ , что и изображение  $D$ , и определенное в той же редкой решетке, что и гиперспектральный снимок  $B$ .

**Шаг 2.** Для каждого пикселя  $\bar{B}_{mn} = (\bar{B}_{mnp}, p = \overline{1, P})$  анализируются соответствующие пиксели изображения  $D$ , входящие в апертуру  $\hat{D}_{mn} = \{D_{rsp} \mid r = \overline{r_1, r_2}, s = \overline{s_1, s_2}, p = \overline{1, P}\}$ . Если  $\Phi(\bar{B}_{mn}, D_{rs}) \geq \tilde{\Phi}$  для всех  $r = \overline{r_1, r_2}$  и  $s = \overline{s_1, s_2}$ , т.е. все пиксели изображения  $D$ , входящие в апертуру  $\hat{D}_{mn}$ , схожи с пикселем  $\bar{B}_{mn}$ , то пиксель  $B_{mn}$  считается «чистым» (т.е. включающим один вполне определенный объект), а соответствующая ему СХ в виде вектора:  $B_{mn} = (B_{mnk}, k = \overline{1, K})$  заносится в список опорных спектральных характеристик  $S = \{S_i \mid i = \overline{1, I}\}$ , где  $S_i = (S_{ik}, k = \overline{1, K})$ ,  $S_{ik} = B_{mnk}$ .

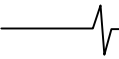
**Шаг 3.** В списке  $S$  удаляются все дублирующиеся спектральные характеристики, т.е.  $S_i$  удаляется, если  $\min_{j=1, I, j \neq i} \Phi(S_i, S_j) \geq \tilde{\Phi}$ . В результате формируется список уникальных опорных спектральных характеристик  $S^* = \{S_i^* \mid i = \overline{1, I^*}\}$ .

**Шаг 4.** Спектральные характеристики  $S_i^*$  из списка  $S^*$  приводятся к спектральному разрешению многозонального изображения  $D$ :

$$\bar{S}_i^* = (\bar{S}_{ip}^*, p = \overline{1, P}), \bar{S}_{ip}^* = \sum_{\lambda=\lambda_{p1}}^{\lambda_{p2}} \delta_p(\lambda) S_{ih(\lambda)}^*, i = \overline{1, I^*}.$$

В результате формируется список  $\bar{S}^* = \{\bar{S}_i^* \mid i = \overline{1, I^*}\}$ .

**Шаг 5.** Для каждого пикселя  $\bar{B}_{mn}$  изображения  $\bar{B}$  анализируются пиксели в апертуре  $\hat{D}_{mn} = \{D_{rsp} \mid r = \overline{r_1, r_2}, s = \overline{s_1, s_2}, p = \overline{1, P}\}$ . Для этого  $D_{rs} = (D_{rsp}, p = \overline{1, P})$  сопоставляется с опорными спектрами  $\bar{S}_i^*$ ,  $i = \overline{1, I^*}$ . В результате для каждой точки  $(r, s)$ ,  $r = \overline{r_1, r_2}$ ,  $s = \overline{s_1, s_2}$ , находится наиболее близкий опорный спектр  $\bar{S}_l^*$  как  $l = \arg \max_i \Phi(\bar{S}_i^*, D_{rs})$ , а также количество точек  $T_i$  в апертуре, для которых найденный спектр является наиболее близким. Процентные доли опорного спектра в апертуре вычисляются как  $\tilde{T}_i = T_i / \omega$ . Найденные опорные спектры и их процентные доли записываются соответственно в списки  $QS_{mn} = \{S_i^* \mid i = \overline{1, \tilde{I}_{mn}}\}$  и  $QT_{mn} = \{\tilde{T}_i \mid i = \overline{1, \tilde{I}_{mn}}\}$ , где  $\tilde{I}_{mn}$  – количество различающихся опорных спектров, найденных в апертуре. Таким образом, для каждой точки  $(m, n)$  изображения  $B$  формируется списки  $QS_{mn}$  и  $QT_{mn}$ , состоящие из спектральных характеристик объектов, образующих «смесь», и их процентных долей в смеси так, что



для каждой точки  $(m, n)$  выполняется соотношение

$$B_{mn} = \sum_{i=1}^{\bar{i}} QT_{mni} (\alpha \cdot QS_{mni} + \beta),$$

где  $\alpha = \{\alpha_k | k = \overline{1, K}\}$ ,  $\beta = \{\beta_k | k = \overline{1, K}\}$  – коэффициенты преобразования опорных СХ, учитывающие различные условия получения опорной СХ, и СХ фактически анализируемой точки. Таким образом, каждой точке изображения  $B'$  ставится в соответствие несколько СХ, образующих эту точку, и их процентные доли в смеси.

### Повышение пространственного разрешения ГСИ

В каждой точке высокодетального снимка восстанавливается спектральная характеристика. Тем самым формируется новое гиперспектральное изображение с пространственным разрешением высокодетального снимка.

Этот процесс включает следующие шаги.

**Шаг 1.** Выполняются шаги 1-4 алгоритма субпиксельного спектрального разделения ГСИ.

**Шаг 2.** Формируется новое изображение  $D^* = \{ D_{rsk}^* | r = \overline{1, R}, s = \overline{1, S}, k = \overline{1, K} \}$  с пространственным разрешением изображения  $D$  и со спектральным разрешением изображения  $B$ . Спектр каждой точки в виде вектора  $D_{rs} = (D_{rsp}, p = \overline{1, P})$  сопоставляется с опорными спектрами  $\overline{S}_i^*$  из списка  $\overline{S}^*$ . В результате сопоставления для произвольной фиксированной точки  $(r, s)$  находится наиболее близкий опорный спектр  $S_l^*$ , где  $l = \arg \max_i \Phi(\overline{S}_i^*, D_{rs})$ . После чего вектор  $D_{rs}^*$  заменяется на  $S_l^*$ . В результате в частой решетке, соответствующей многозональному изображению  $D$ , формируется гиперспектральное изображение  $D^*$ .

**Шаг 3.** Исходя из физического принципа съемки:

$$D_{rsp} = \sum_{\lambda=\lambda_{p1}}^{\lambda_{p2}} \delta_p(\lambda) D_{rsk(\lambda)}^*,$$

т.е. сумма элементов полученного спектра  $D_{rs}^*$  в диапазонах съемки многозонального изображения должна быть равна значению яркости в соответствующем пикселе  $D_{rs}$ . Однако из-за различных условий получения опорной СХ (на основе которой сформирован спектр точки  $D_{rs}^*$ ) и фактически анализируемой точки  $(r, s)$  это равенство не выполняется. Для разрешения этой си-

туации рассчитываются мультипликативные коэффициенты коррекции  $\gamma_{rsk}$ , такие что:

$$D_{rsp} = \sum_{\lambda=\lambda_{p1}}^{\lambda_{p2}} \delta_p(\lambda) \gamma_{rsk(\lambda)} D_{rsk(\lambda)}^*,$$

где  $\gamma_{rsk(\lambda)}$  – непрерывная функция от  $k(\lambda)$ , построенная путем линейной интерполяции  $P$ -точек:

$$\gamma_{rsp} = D_{rsp} / \overline{D}_{rsp}^*, \quad \overline{D}_{rsp}^* = \sum_{\lambda=\lambda_{p1}}^{\lambda_{p2}} \delta_p(\lambda) D_{rsk(\lambda)}^*, \quad p = \overline{1, P}.$$

Результирующее изображение  $D^{**}$  находится как

$D_{rsk}^{**} = \gamma_{rs}(k) D_{rsk}^*$ . В результате в узлах частой решетки, соответствующей высокодетальному многозональному изображению  $D$ , формируется итоговое гиперспектральное изображение  $D^{**}$ , обладающее высокими и пространственным и спектральным разрешениями.

### Эксперименты и выводы

Испытания рассмотренных алгоритмов выполнялись следующим образом (рис. 1). В качестве опорной информации использовался авиационный гиперспектральный снимок высокого пространственного разрешения  $B^*$ , состоящий из  $K$ -каналов. Этот снимок использовался в качестве эталонного. На основе него путем операции усреднения и прореживания в  $q$  раз строилось гиперспектральное изображение низкого пространственного разрешения  $B$ . Путем объединения спектральных каналов формировался многозональный снимок  $D$ , включающий  $L$ -диапазонов, где  $L \ll K$ . Изображение  $D$  обладало пространственным разрешением исходного изображения  $B^*$  с существенно меньшим спектральным разрешением. Выполнялось повышение пространственного разрешения изображения  $B$  на основе спектрального разделения (в качестве меры сходства  $\Phi$  использовалась мера спектрального угла [10]). В результате формировалось изображение  $D^{**}$ . Качество работы алгоритма спектрального разделения оценивалось путем сопоставления СХ одноименных элементов изображений  $D^{**}$  и  $B^*$ . С использованием меры спектрального угла оценивалась ошибка спектрального разделения. За ошибочные принимались точки, для которых эта мера превышала 5 градусов. В результате на достаточно представительном статистическом материале (при  $q=22$  и  $L=4$ ) установлено, что доля ошибочных точек составила 4%, средняя величина этой ошибки по всем точкам составляет 1.3 градуса, а СКО – 1.7 градуса.

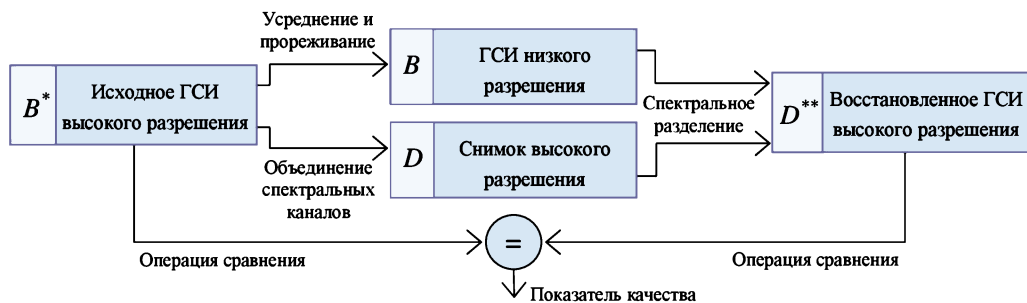


Рис. 1. Схема расчета показателя качества повышения разрешения ГСИ



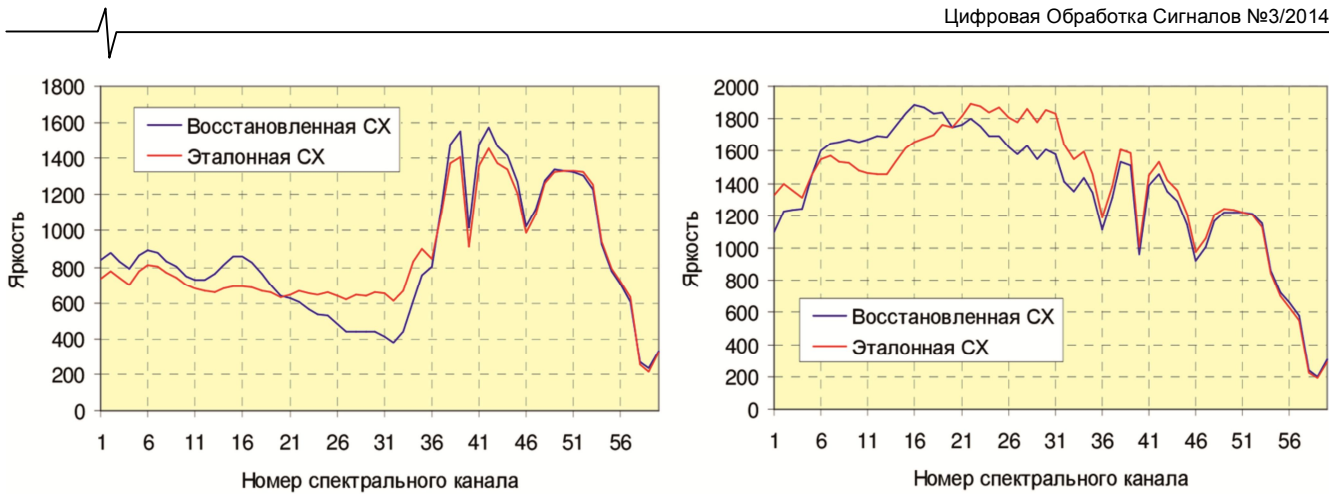


Рис. 2. Эталонная и восстановленная СХ

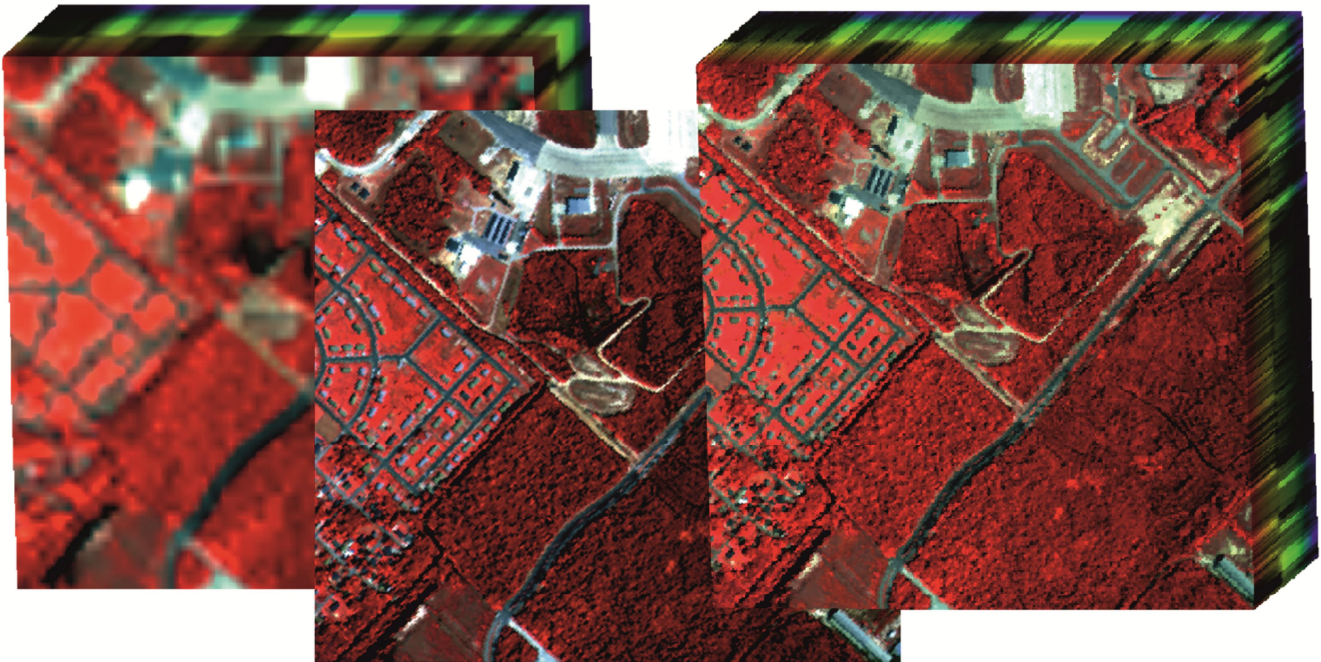


Рис. 3. Комплексование ГСИ с данными высокого пространственного разрешения

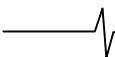
На рис. 2 слева приведены две СХ, построенные для вполне определенного объекта: красным цветом – для ГСИ высокого разрешения  $B^*$ , синим цветом – для восстановленного ГСИ  $D^{**}$ . На рис. 2 справа представлены аналогичные результаты для другого объекта. Как видим, восстановленные СХ достаточно близки к эталонным.

В качестве примера на рис. 3 представлен результат комплексования: слева – ГСИ с низким пространственным разрешением; в центре – многозональный снимок высокого пространственного разрешения; справа – результирующее ГСИ.

Таким образом, предложен подход к спектральному разделению гиперспектральных изображений с привлечением в качестве опорной информации высокдетальных многозональных снимков. Разработан алгоритм повышения пространственного разрешения гиперспектральных снимков, основанный на спектральном разделении пикселей ГСИ. Выполнена оценка качества алгоритма спектрального разделения и повышения пространственного разрешения ГСИ на основе натурной и модельной информации.

### Литература

1. Еремеев В.В. «Современные направления работ по анализу и повышению качества космических изображений поверхности Земли» // Цифровая обработка сигналов. №1. 2012. С. 38 – 44.
2. Ахметов Р.Н., Стратилатов Н.Р., Юдаков А.А., Везенов В.И., Еремеев В.В. «Модели формирования и некоторые алгоритмы обработки гиперспектральных изображений» // Исследование Земли из космоса. №5. 2012. С. 17-28.
3. Lucas Parra, Clay Spence, Paul Sajda, Andreas Ziehe, Klaus-Robert Müller, «Unmixing Hyperspectral Data», in Advances in Neural Information Processing 12 (Proc. NIPS'99). 2000. PP. 942-948.
4. J.J. Settle, «Linear mixing and the estimation of endmembers for the spectral decomposition of remotely sensed scenes», SPIE Remote Sensing for Geology, 2960. 1996. PP. 104-109.
5. Lordache, M.-D.; Bioucas-Dias, J.M.; Plaza, A., «Sparse Unmixing of Hyperspectral Data», Geoscience and Remote Sensing, IEEE Transactions on , vol.49, no.6. 2011. PP. 2014-2039.



6. Антонушкина С.В., Еремеев В.В., Макаренков А.А., Москвитин А.Э. «Особенности анализа и обработки информации от систем гиперспектральной съемки земной поверхности» // Цифровая обработка сигналов. №4. 2010. С. 38-43.

7. Еремеев В.В., Макаренков А.А., Москвитин А.Э., Юдаков А.А. «Повышение четкости отображения объектов на данных гиперспектральной съемки земной поверхности» // Цифровая обработка сигналов. №3. 2012. С. 35 – 40.

8. Еремеев В.В. «Современные проблемы обработки данных дистанционного зондирования Земли» // Радиотехника. №3. 2012. С. 54-64.

9. Еремеев В.В., Макаренков А.А., Москвитин А.Э., Мятлов Г.Н. «Повышение информативности материалов съемки Земли за счет объединения гиперспектральной информации с данными от других съемочных систем» // Цифровая обработка сигналов. №1. 2012. С. 38 – 44.

10. Yugas, R.H., Goetz, A. F. H., and Boardman, J. W.,

«Discrimination among semiarid landscape endmembers using the spectral angle mapper (SAM) algorithm», In Summaries of the Third Annual JPL Airborne Geoscience Workshop, JPL Publication 92-14, vol. 1. 1992. PP.147-149.

### **INCREASING OF SPATIAL RESOLUTION OF HYPERSPECTRAL IMAGERY OF EARTH BASED ON FUSION WITH HIGH RESOLUTION IMAGES**

*Eremeev V.V., Makarenkov A.A., Moskvitin A.E.*

Increasing of spatial resolution of hyperspectral data were reviewed. A new approach to solving this problem is proposed. It is based on spectral unmixing algorithm. Spectral unmixing is done by incorporating of the synchronous high resolution multispectral survey data. Results of processing of the actual hyperspectral data were presented.

## **Журнал «Цифровая обработка сигналов»**

**покоряет вершины гор ...**



**Пик Курсанты**



**Эльбрус**

УДК 004.932

## АЛГОРИТМ ПРЕДВАРИТЕЛЬНОГО ОЦЕНИВАНИЯ ПРОСТРАНСТВЕННОЙ ОРИЕНТАЦИИ ОБЪЕКТА С ПОМОЩЬЮ ДЕСКРИПТОРА ВНЕШНЕГО КОНТУРА

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

*Бабаян П.В., к.т.н., доцент кафедры автоматики и информационных технологий в управлении РГРТУ, e-mail: aitu@rsreu.ru;*

*Смирнов С.А., н.с. кафедры автоматики и информационных технологий в управлении РГРТУ;*

*Масленников Е.А., студент РГРТУ.*

**Ключевые слова:** определение ориентации, дескриптор внешнего контура, геосфера, углы Эйлера.

### Введение

Оценивание пространственной ориентации объекта по его изображению является актуальной задачей технического зрения [1]. Оценивание пространственной ориентации объектов наиболее востребовано в таких приложениях как:

- управление наземными мобильными роботами, пилотируемыми и беспилотными летательными аппаратами;
- автоматическая дозаправка летательных аппаратов;
- стыковка космических кораблей [2];
- формирование изображений в системах улучшенного и синтезированного видения;
- анализ и индексация видеосюжетов.

В работах [3, 4] был предложен подход к решению задачи оценивания пространственной ориентации объекта, состоящий из двух этапов: обучение и оценивание. На этапе обучения происходит расчет дескрипторов эталонных изображений объекта при равномерно распределенных на сфере ориентациях. Эталонные изображения формируются на основе 3D-модели исследуемого объекта. На этапе оценивания для определения ориентации объекта находится дескриптор эталонного изображения, который в наибольшей степени соответствует дескриптору наблюдаемого изображения. Представлено 3 алгоритма оценивания пространственной ориентации объекта, основанных на данном подходе: с использованием БПФ, с использованием текстурных характеристик, с использованием структурного анализа. Следует отметить, что предложенные алгоритмы обладают рядом недостатков. Так, алгоритм на основе текстурного анализа чувствителен к шуму. Два других алгоритма обладают большой вычислительной сложностью. В настоящей работе представлен алгоритм, при разработке которого предпринималась попытка удовлетворить требованиям к невысокой вычислительной сложности и устойчивости к шуму.

Основным отличием предлагаемого алгоритма от разработанных в [4], является способ описания изображения объекта. В качестве описания изображения используется дескриптор внешнего контура. Дескриптор внешнего контура обладает следующими преимуществами:

- независимость от текстурных характеристик изобра-

*Рассмотрена задача оценивания пространственной ориентации объекта по его двумерному изображению с использованием известной трехмерной модели объекта. Предложен алгоритм предварительного оценивания пространственной ориентации объекта с помощью дескриптора внешнего контура. Представлены результаты экспериментальных исследований предложенного алгоритма.*

жения объекта;

- устойчивость к изменению освещения объекта;
- небольшой размер дескриптора;
- низкие вычислительные затраты для вычисления дескриптора и функции сопоставления дескриптора;
- высокая устойчивость к шуму.

Результатом работы представленного в статье алгоритма является множество ориентаций-кандидатов, в наилучшей степени соответствующих наблюдаемому изображению объекта по критерию близости дескриптора внешнего контура. Выбор из этого множества одного элемента, который принимается за оценку текущей ориентации объекта, не является предметом рассмотрения в настоящей статье и может осуществляться на основе анализа яркостных, геометрических, спектральных параметров объекта, а также их поведения во времени.

### Представление ориентации объекта в пространстве. Дискретизация сферы

Для математического описания пространственной ориентации объекта будем использовать углы Эйлера [5, 6]  $(\alpha, \beta, \gamma)$ . Для нашей задачи примем соглашение, при котором осуществляются последовательные повороты вокруг осей  $X, Y, Z$  в глобальной системе координат. Углы Эйлера в данном случае будут соответствовать матрице вращения, которую можно вычислить по формуле:

$$R = R_z \cdot R_y \cdot R_x, \quad (1)$$

где  $R_z, R_y, R_x$  – матрицы вращения вокруг осей  $Z, Y, X$  соответственно.

Отметим, что в этом случае третий угол Эйлера  $\gamma$  соответствует углу поворота камеры вокруг оптической оси.

Поскольку в предложенном алгоритме осуществляется сопоставление дескриптора наблюдаемого с эталонными дескрипторами, возникает задача построения множества эталонов, равномерно охватывающих все возможные ориентации объекта интереса. Использо-

ние равномерно дискретизированных углов Эйлера не позволяет построить такое множество, поэтому для решения этой задачи в данной работе был использован алгоритм построения геосферы [7].

Для описания равномерно распределенных ориентаций рассмотрим сферу единичного радиуса, в центре которой находится объект интереса. Каждой точке сферы  $(x, y, z)$  можно сопоставить углы Эйлера  $\alpha$  и  $\beta$ , оставив  $\gamma$  свободным (поворот камеры вокруг оптической оси). Пары углов  $(\alpha, \beta)$ , равномерно распределенные по сфере, получаются в результате построения геосферы фактора  $g$ , вписанной в рассматриваемую единичную сферу. Вершины построенной геосферы соответствуют первым двум углам Эйлера. Параметр  $\gamma$  не дискретизируется, поскольку используемая в предложенном алгоритме критериальная функция, рассматриваемая ниже, инвариантна к параметру  $\gamma$ .

Выбор фактора геосферы влияет на количество точек геосферы ( $N_g$ ) и максимально возможным углом между радиус-векторами точки единичной сферы и ближайшей точки геосферы ( $E_{\max}$ ). В табл. 1 приведены параметры геосферы в зависимости от фактора геосферы.

Таблица 1. Параметры геосферы в зависимости от фактора геосферы

$g$	$N_g$	$E_{\max}, ^\circ$
1	12	17,6
2	42	10,2
3	92	6,8
4	162	5,4
5	252	4,3
6	362	3,6
7	492	3,1
8	642	2,7
9	812	2,4

Выбор фактора геосферы должен происходить исходя из компромисса между вычислительной сложностью, пропорциональной параметру  $N_g$  и требуемой точностью оценивания ориентации, который имеет один порядок с параметром  $E_{\max}$ . В данной работе для построения геосферы был выбран фактор  $g = 6$ .

Экспериментальные исследования в настоящей работе проводились на синтезированных изображениях с использованием 3D-модели пассажирского самолёта Airbus A320.

### Дескриптор внешнего контура

Для каждой точки геосферы производится синтез изображения объекта интереса и вычисляется дескриптор, основанный на характеристиках контура объекта. После бинаризации эталонного изображения (рис. 1а) получаем бинарное изображение объекта с ярко выраженным контуром (рис. 1б). Данное изображение переводится в полярную систему координат с началом координат в центре бинарного изображения объекта (рис. 1в). Зададимся длиной дескриптора  $N_D$ . Для вычисления дескриптора объекта формируем функцию

$$\rho_i = \tilde{D}(\phi_i), \text{ выбирая для каждого } \phi_i = \frac{(i-1) \cdot 2\pi}{N_D},$$

$i = \overline{1, N_D}$ , максимальное соответствующее ему  $\rho$ . Для удаления артефактов, связанных с синтезом изображения, функция  $\tilde{D}(\phi_i)$  подвергается медианной фильтрации с окном фильтра длиной 5 элементов. Полученную дискретную функцию  $\rho_i = D(\phi_i)$  (рис. 1г) назовём дескриптором внешнего контура объекта. Для упрощения дальнейшего изложения будем использовать аргумент  $i$  вместо  $\phi_i$ , таким образом  $\rho_i = D(i)$ .

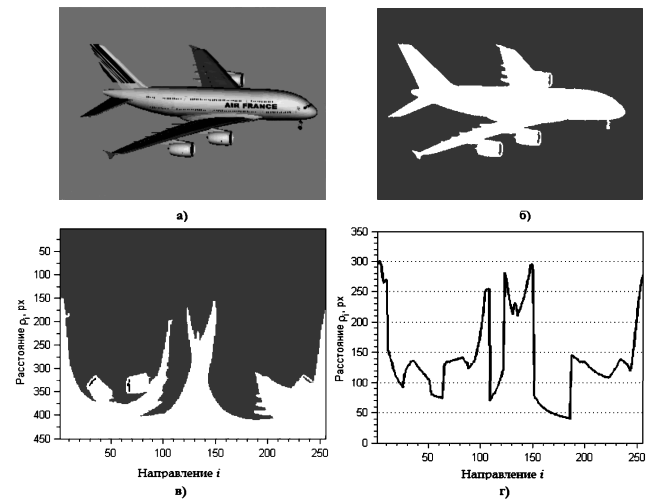


Рис. 1. Этапы вычисления дескриптора внешнего контура: а – исходное изображение объекта; б – бинарное изображение объекта; в – бинарное изображение в полярной системе координат; г – дескриптор внешнего контура объекта

Экспериментально было подтверждено, что данный дескриптор обладает высокой устойчивостью к шуму. Для оценки устойчивости к шуму был введен показатель отличия дескрипторов, вычисляющийся по формуле:

$$\varepsilon_\sigma = \frac{1}{N_D} \sum_{i=1}^{N_D} \frac{|D(i) - G(i)|}{D(i)}, \quad (2)$$

где  $D(i)$  –  $i$ -й элемент дескриптора без зашумления,  $G(i)$  –  $i$ -й элемент дескриптора, вычисленного на зашумленном изображении с ОСШ  $= \sigma$ .

Величина  $\varepsilon_\sigma$  описывает, насколько изменяется дескриптор зашумленного изображения объекта относительно дескриптора того же изображения объекта без зашумления. Значения показателя  $\varepsilon_\sigma$  для различного уровня зашумления приведены в табл. 2.

Таблица 2. Показатель отличия дескриптора на зашумленном изображении

ОСШ	50	20	10	5
$\varepsilon_\sigma, \%$	0,03	0,16	0,65	2,63

### Сопоставление дескрипторов

Определение ориентации объекта происходит в результате сопоставления дескриптора наблюдаемого изображения с каждым дескриптором эталонной геосферы. Определение сходства дескриптора наблюдаемого изо-



бражения с дескрипторами эталонных изображений осуществляется с использованием критериальной функции:

$$f_{crit}(j) = \min_{s \in S} \left[ \sum_{i=1}^{N_D} (D_0(i-s) - D_j(i))^2 \right], \quad (3)$$

$$S = \{s : s = \overline{1, N_D}\}$$

где  $D_0$  – дескриптор наблюдаемого изображения,  $D_j$  – дескриптор текущего эталонного изображения,  $i$  – индекс элемента дескриптора,  $j$  – индекс эталонного дескриптора,  $s$  – величина циклического сдвига дескриптора.

Выборный тип критериальной функции обеспечивает инвариантность дескриптора к повороту изображения.

Индекс  $j$  эталонного дескриптора соответствует точке геосферы, то есть двум углам  $\alpha$  и  $\beta$  из тройки углов Эйлера.

Обозначим через  $s_0$  значение сдвига  $s$ , при котором выражение в квадратных скобках в (3) принимает минимальное значение.

$$s_0(j) = \arg \min_{s \in S} \left[ \sum_{i=1}^{N_D} (D_0(i-s) - D_j(i))^2 \right], \quad (4)$$

$$S = \{s : s = \overline{1, N_D}\}$$

Отсюда угол  $\gamma$  вычисляется по формуле:

$$\gamma = \frac{s_0 \cdot 2\pi}{N_D} \quad (5)$$

В результате вычисления критериальной функции для каждого дескриптора из эталонной базы получается вектор значений критериальной функции:

$$M = (f_{crit}(j) | j = \overline{1, N_g}) \quad (6)$$

Обозначим через  $M_{sort}$  вектор значений критериальной функции, сортированный по возрастанию. Были проведены исследования, в результате которых было обнаружено, что, как правило, хотя бы один из первых  $m = N_g \cdot 0,025$  элементов вектора  $M_{sort}$  соответствует текущей ориентации объекта с ошибкой оценивания ориентации до  $10^\circ$ . Так, для геосферы фактора 6, получаем  $m = 9$  ориентаций-кандидатов.

Среди полученных  $m$  ориентаций-кандидатов встречаются похожие ориентации. Для сокращения неопределенности оценивания ориентации производится объединение групп ориентаций-кандидатов в кластеры [8]. Для кластеризации строится граф, вершинами которого являются ориентации-кандидаты. Между вершинами строится дуга, только в случае, когда угловое отличие между ними не превышает заданного порога. Связные компоненты графа образуют кластеры. В качестве углового отличия между ориентациями-кандидатами используется расстояние в гиперкомплексном пространстве между соответствующими им кватернионами.

На рис. 2 показаны ориентации, отобранные как кандидаты при оценивании ориентации  $E = [-80,3^\circ -49^\circ 83,4^\circ]$ .

Рассчитанные кластеры для приведенного выше примера показаны в табл. 3

Таблица 3. Результат кластеризации

№ кластера	1	2	3	4
Элементы кластера	1 3 4 5	2 6 9	7	8

Элементом, представляющим кластер, можно считать элемент кластера соответствующий наименьшему значению критериальной функции или усредненное значение всех элементов в кластере. На рис. 3 показаны ориентации-кандидаты после кластеризации.

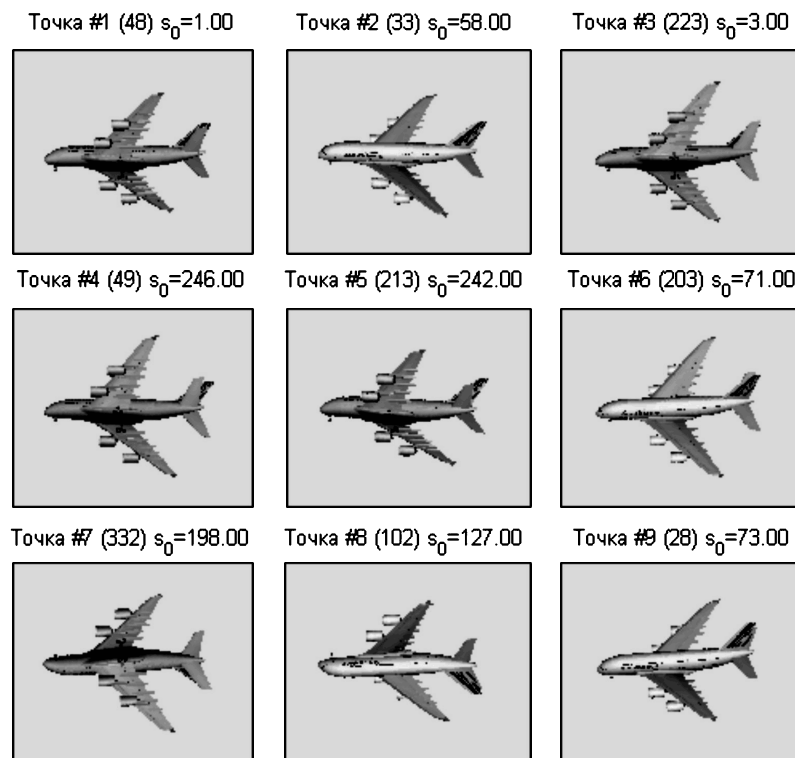


Рис. 2. Изображения 9 ориентаций-кандидатов

Кластер #1 (Точка=48)  $s_0=1$ Кластер #2 (Точка=33)  $s_0=58$ Кластер #3 (Точка=223)  $s_0=3$ Кластер #4 (Точка=49)  $s_0=246$ 

Рис. 3. Изображения 4 лучших кандидатов после этапа кластеризации

### Экспериментальные исследования

Для проведения экспериментальных исследований было синтезировано 500 изображений. Распределение ориентаций объекта было близко к равномерному. Полученные изображения были искажены равномерным аддитивным шумом с различной интенсивностью. На стадии обучения были рассчитаны дескрипторы объекта по геосфере фактора 6. В ходе исследования производилась совмещение дескриптора наблюдаемого изображения с эталонными дескрипторами. Затем проводилась кластеризация наилучших отобранных точек.

Частота события, заключающегося в том, что ориентация, соответствующая первому элементу вектора  $M_{sort}$ , является правильным результатом оценивания для различных уровней зашумления показана в табл. 4.

Таблица 4

ОСШ	$\infty$	25	15	10	5
Частота	98,5%	98,1%	97,2%	94,5%	77,9%

Была рассчитана величина  $m$  – минимальное количество элементов вектора  $M_{sort}$ , среди которых присутствует элемент, соответствующий значению ориентации с допустимой ошибкой  $e$ . Значения  $m$  для различных уровней зашумления и величины допустимой ошибки  $e < 10^\circ$  показаны в табл. 5. Вычисленные значения становятся полезными на этапе кластеризации отобранных дескрипторов.

Таблица 5

ОСШ	$\infty$	25	15	10	5
Величина $m$ при $e < 10^\circ$	4	3	5	9	30

В ходе исследования было выявлено, что в результате кластеризации ориентаций-кандидатов всегда получалось от 1 до 5 кластеров. Распределение частот различного количества кластеров показано в табл. 6.

Таблица 6

Количество кластеров	1	2	3	4	5
Частота	3,2%	53,5%	32%	9,5%	1,8%

### Заключение

Разработанный алгоритм предварительного определения ориентации подтвердил свою работоспособность в ходе экспериментальных исследований. Данный алгоритм обладает невысокой вычислительной сложностью (составление дескрипторов производится в 20 раз быстрее, чем в алгоритмах, предложенных в [4]). Экспериментальные исследования подтвердили устойчивость алгоритму к шуму на наблюдаемом изображении.

В отличие от алгоритмов, предложенных в [4], результатом рассматриваемого в этой работе алгоритма является набор ориентаций-кандидатов. Следуя данным из табл. 4, можно сделать вывод, что если считать результатом оценивания первую ориентацию-кандидата, то вероятность правильного оценивания будет составлять от 78% до 98,5% в зависимости от интенсивности зашумления изображения.

Работа выполнена при поддержке гранта RFMEFI57414X0056 Министерства образования и науки РФ.

### Литература

- 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
- Бахшиев А.В., Степанов Д.Н., Кирпань Н.А. Программный комплекс определения пространственной ориентации объектов по телевизионному изображению // Робототехника и техническая кибернетика. – 2013. – №1 (1) – С.71-75
- Бабаян П.В., Масленников Е.А. Алгоритмы оценивания пространственной ориентации объекта по его двумерному изображению // Доклады 15-й междунар. конф. DSPA-2013 том-2. – М., 2013 – С.58-62.
- Алпатов Б.А., Бабаян П.В., Масленников Е.А. Алгоритмы оценивания ориентации объекта по его двумерному изображению в бортовых системах видеослежения // Вестник Рязанского государственного радиотехнического университета. – Рязань, 2013. – №3. – С. 3-8.
- Bekir E. Introduction to Modern Navigation Systems // World Scientific Publishing Co. Pte. Ltd. 255 p., 2007.
- Slabaugh, G.G. (1999). Computing Euler angles from a rotation matrix. Retrieved on August, 6, 2000.
- Saff E., Kuijlaars A. «Distributing many points on a sphere» // The Mathematical Intelligencer, 1997, Vol. 19, No. 1. – pp. 5-11.
- Schaeffer S. E. Graph clustering // Computer Science Review. – 2007. – Т. 1. – №. 1. – С. 27-64.

### ALGORITHM OF PRELIMINARY POSE ESTIMATION USING EXTERNAL CONTOUR DESCRIPTOR

*Alpatov B.A., Babayan P.V., Maslennikov E.A.*

Problem of object three-dimensional pose estimation from its two-dimensional image using 3d-model of object is considered. The algorithm of preliminary pose estimation using external contour descriptor is offered. Results of experimental research of these algorithms are shown.

УДК. 621.391

## МЕТОД ВЫДЕЛЕНИЯ ДВИЖУЩИХСЯ ОБЪЕКТОВ В ВИДЕОПОТОКЕ И ОЦЕНКА ТОЧНОСТИ ОПРЕДЕЛЕНИЯ ИХ КООРДИНАТ

*Медведева Е.В., д.т.н., профессор кафедры радиоэлектронных средств Вятского государственного университета, e-mail: emedv@mail.ru;*

*Карлушин К.А., аспирант Вятского государственного университета, e-mail: konstantin.karlushin@gmail.com.*

**Ключевые слова:** выделение движущихся объектов, оценка точности определения координат, изображения, многомерные цепи Маркова.

### Введение

В некоторых системах видеонаблюдения необходимо выделить движущиеся объекты на сравнительно неподвижном фоне. Как правило, такие системы должны с большой вероятностью обнаруживать объекты интереса. При этом диапазон варьирования размеров объектов интереса может быть большим, а количество движущихся объектов – априорно неизвестным. Кроме того, обработка данных должна проводиться в реальном масштабе времени, что неизбежно накладывает ограничения на сложность используемых алгоритмов.

Для решения задачи выделения объектов существует много различных методов (разностный, корреляционный, на основе байесовской сегментации и др.) [1–4], но одни из них могут приводить к существенным погрешностям в оценке координат, а другие требуют относительно большого объема вычислительных ресурсов, особенно, когда в видеокадре необходимо выделить несколько объектов.

Поэтому разработка эффективных методов выделения объектов движения в видеопотоке, позволяющих в реальном масштабе времени выделять объекты интереса с высокой точностью, является актуальной задачей.

Поскольку последовательности видеокадров представляют собой трехмерные сигналы, для выделения движущихся объектов необходимо использовать пространственно-временную математическую модель изображения. Пространственно-временная модель должна быть адекватна реальным последовательностям изображений и обладать высокой вычислительной эффективностью, чтобы алгоритмы обработки изображений, синтезируемые на ее основе, требовали небольших вычислительных затрат.

Такой математической моделью является модель на основе трехмерных цепей Маркова.

Предполагается, что цифровые полутоновые изображения (ЦПИ) в видеопоследовательности представлены  $g$ -разрядными двоичными числами. Это позволяет разбить последовательность ЦПИ на последовательности  $g$  разрядных двоичных изображений (РДИ), каждое из которых является суперпозицией трех простых цепей Маркова с двумя равновероятными дискретными состояниями  $M_1^{(l)}$  и  $M_2^{(l)}$  и матрицами вероятностей переходов по горизонтали

${}^1\Pi = \|\pi_{ij}^{(l)}\|_{2 \times 2}$ , вертикали  ${}^2\Pi = \|\pi_{ij}^{(l)}\|_{2 \times 2}$  и времени

Предложен метод выделения движущихся объектов в видеопотоке. Метод основан на представлении последовательности изображений трехмерной цепью Маркова и выделении контуров движущихся объектов по вычисленной величине количества информации в элементах изображения. Проведена оценка точности определения координат движущихся объектов. Разработанный метод требует небольших вычислительных ресурсов, что делает возможным его применение при обработке данных в реальном масштабе времени.

$${}^4\Pi = \|\pi_{ij}^{(l)}\|_{2 \times 2} \quad (l = \overline{1, g}, i, j = \overline{1, 2}).$$

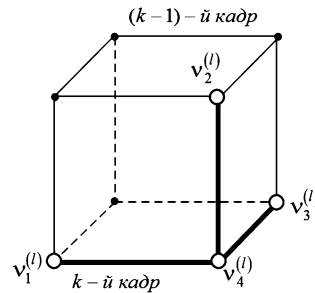


Рис. 1. Фрагмент последовательности РДИ

На приведенном фрагменте последовательности РДИ (рис. 1), элемент изображения  $v_4^{(l)}$  зависит от состояний трех окрестных элементов  $\Lambda_{i,j,k}^{(l)} = \{v_1^{(l)}, v_2^{(l)}, v_3^{(l)}\}$ , расположенных в  $k$ -м и  $(k-1)$ -м кадрах.

Количество информации в элементе  $v_4^{(l)}$  относительно состояний элементов окрестности  $\Lambda_{i,j,k}^{(l)} = \{v_1^{(l)}, v_2^{(l)}, v_3^{(l)}\}$  можно определить как разность собственной информации в элементе  $v_4^{(l)}$  и взаимной информации, полученной от элементов окрестности  $\Lambda_{i,j,k}^{(l)}$  [5–7]:

$$I(v_4^{(l)} | v_1^{(l)}, v_2^{(l)}, v_3^{(l)}) = I(v_4^{(l)}) - I(v_1^{(l)}, v_2^{(l)}, v_3^{(l)}, v_4^{(l)}) = - \left[ \log p(v_4^{(l)}) + \log \frac{\prod_4 p(v_i^{(l)}, v_j^{(l)}) p(v_1^{(l)}, v_2^{(l)}, v_4^{(l)}, v_4^{(l)})}{\prod_{i=1}^4 p(v_i^{(l)}) \prod p(v_i^{(l)}, v_j^{(l)}, v_k^{(l)})} \right] = - \log \frac{w(v_4^{(l)} | v_1^{(l)}) w(v_4^{(l)} | v_2^{(l)}) w(v_4^{(l)} | v_3^{(l)}) w(v_4^{(l)} | v_1^{(l)}, v_2^{(l)}, v_3^{(l)})}{w(v_4^{(l)} | v_1^{(l)}, v_2^{(l)}) w(v_4^{(l)} | v_1^{(l)}, v_3^{(l)}) w(v_4^{(l)} | v_2^{(l)}, v_3^{(l)})}, \quad (1)$$

где произведения, обозначенные символом  $\prod$  вычисляются по всем возможным несовпадающим комбинациям различных подстрочных индексов трехмерного слу-

чайного поля;  $p(v_1^{(l)}, v_2^{(l)}, v_4^{(l)}, v_4^{(l)})$ ,  $p(v_i^{(l)}, v_j^{(l)}, v_k^{(l)})$ ,  $p(v_i^{(l)}, v_j^{(l)})$ ,  $i = j = k = \overline{1,4}$ ;  $i \neq j \neq k$  - совместные плотности вероятностей значений элементов;  $p(v_i^{(l)})$ ,  $i = \overline{1,4}$  - априорные плотности вероятностей значений элементов;  $w(v_4^{(l)} | v_i^{(l)})$ ,  $i = \overline{1,3}$  - одномерные плотности вероятностей переходов;  $w(v_4^{(l)} | v_i^{(l)}, v_j^{(l)})$ ,  $i = j = \overline{1,3}$ ;  $i \neq j$  - плотности вероятностей переходов в двумерной цепи Маркова;  $w(v_4^{(l)} | v_1^{(l)}, v_2^{(l)}, v_3^{(l)})$  - плотность вероятности переходов в трехмерной цепи Маркова.

Плотность вероятности переходов в трехмерной двочной цепи Маркова  $w(v_4^{(l)} | \Lambda_{i,j,k}^{(l)})$  можно представить в форме

$$w(v_4^{(l)} | \Lambda_{i,j,k}^{(l)}) = \sum_{i,r=1}^2 \pi \left( v_4^{(l)} = M_i^{(l)} | v_1^{(l)} = M_j^{(l)}; v_2^{(l)} = M_k^{(l)}; v_3^{(l)} = M_r^{(l)} \right) \times \delta(v_1^{(l)} - M_j^{(l)}) \delta(v_2^{(l)} - M_k^{(l)}) \delta(v_3^{(l)} - M_r^{(l)}), \quad (2)$$

где  $\delta(\cdot)$  - дельта функция.

В трехмерной математической модели вероятности смены состояний элемента  $v_4^{(l)}$  зависят от различных сочетаний состояний элементов окрестности  $\Lambda_{i,j,k}^{(l)}$  и их корреляции с элементом  $v_4^{(l)}$ . Все возможные значения этих вероятностей для различных комбинаций состояний элементов окрестности  $\Lambda_{i,j,k}^{(l)}$  могут быть вычислены в соответствии с аргументом выражения (1) и сведены в матрицу  $\Pi$  вида:

$$\Pi = \begin{pmatrix} \pi_{iii}^{(l)} & \pi_{ijj}^{(l)} \\ \pi_{iji}^{(l)} & \pi_{ijj}^{(l)} \\ \vdots & \vdots \\ \pi_{jji}^{(l)} & \pi_{jjj}^{(l)} \end{pmatrix} = \begin{pmatrix} \alpha_1^{(l)} & \alpha_1^{(l)} \\ \alpha_2^{(l)} & \alpha_2^{(l)} \\ \vdots & \vdots \\ \alpha_8^{(l)} & \alpha_8^{(l)} \end{pmatrix}, \quad i, j = \overline{1,2}; \quad i \neq j, \quad (3)$$

элементы которой симметричны и удовлетворяют условию нормировки  $\alpha_q + \alpha_q^{(l)} = 1$ .

Например, выражения для вычисления элементов первой строки матрицы  $\Pi$ , имеют вид:

$$\alpha_1^{(l)} = \pi_{iii}^{(l)} = \frac{1 \pi_{ii}^{(l)} \cdot 2 \pi_{ii}^{(l)} \cdot 4 \pi_{ii}^{(l)} \cdot 7 \pi_{ii}^{(l)}}{3 \pi_{ii}^{(l)} \cdot 5 \pi_{ii}^{(l)} \cdot 6 \pi_{ii}^{(l)}}, \quad (4)$$

$$\alpha_1^{(l)} = \pi_{ijj}^{(l)} = \frac{1 \pi_{ij}^{(l)} \cdot 2 \pi_{ij}^{(l)} \cdot 4 \pi_{ij}^{(l)} \cdot 7 \pi_{ii}^{(l)}}{3 \pi_{ii}^{(l)} \cdot 5 \pi_{ii}^{(l)} \cdot 6 \pi_{ii}^{(l)}},$$

где  ${}^r \pi_{ii}^{(l)}$  ( $r = \overline{1,7}$ ) - элементы матриц вероятностей переходов в одномерных цепях Маркова с двумя состояниями по горизонтали  ${}^1 \Pi^{(l)}$ , вертикали  ${}^2 \Pi^{(l)}$ , времени  ${}^4 \Pi^{(l)}$  и четырех сопутствующих матриц вероятностей переходов:

$$\begin{aligned} {}^3 \Pi^{(l)} &= {}^1 \Pi^{(l)} \times {}^2 \Pi^{(l)}; \\ {}^5 \Pi^{(l)} &= {}^1 \Pi^{(l)} \times {}^4 \Pi^{(l)}; \\ {}^6 \Pi^{(l)} &= {}^2 \Pi^{(l)} \times {}^4 \Pi^{(l)}; \\ {}^7 \Pi^{(l)} &= {}^3 \Pi^{(l)} \times {}^4 \Pi^{(l)}. \end{aligned} \quad (5)$$

Вычисление остальных элементов матрицы  $\Pi^{(l)}$  осуществляется в соответствии с состоянием элементов окрестности  $\Lambda_{i,j,k}^{(l)}$ .

На рис. 2 показаны зависимости вероятностей переходов между состояниями элементов в последовательности РДИ по горизонтали, вертикали и между кадрами от номера разряда ЦПИ. Графики получены усреднением статистических характеристик по большому количеству реальных изображений, полученных стационарной видеокамерой, аналогичных рис. 3.



Рис. 2. Зависимости вероятностей переходов между состояниями элементов в последовательности РДИ (8 - старший разряд; 1 - младший разряд)

Из приведенного графика следует, что наибольшей корреляцией обладают элементы, расположенные в старших разрядах ЦПИ. Поэтому основные детальные области можно выделить по РДИ старших разрядов ЦПИ. А РДИ младших разрядов (при  $\pi_{ii} \rightarrow 0,5$ ) будут составлять фон изображения в виде двумерного шума.

Для выделения движущихся объектов в каждом элементе РДИ в соответствии с состояниями элементов окрестности вычисляют величину количества информации (1). Для видеопоследовательностей, полученных стационарной видеокамерой при неизменном уровне освещенности, будет характерно незначительное изменение вероятностей переходов между элементами изображения от кадра к кадру в течение длительного промежутка времени. Поэтому предполагалось, что вероятности переходов между элементами видеопоследовательности постоянны и априорно известны, причем корреляция элементов между кадрами ( ${}^4 \pi_{ii}$ ) больше, чем внутри кадра ( ${}^1 \pi_{ii}$ ,  ${}^2 \pi_{ii}$ ) (рис. 2).

При отсутствии объекта (контрастной области), состояния окрестных элементов  $\Lambda_{i,j,k}^{(l)}$  будут совпадать с  $v_4^{(l)}$ , и количество информации в элементе  $v_4^{(l)}$  будет минимально.

При наличии неподвижного объекта, на его границе, состояние элемента  $v_4^{(l)}$  будет совпадать с состоянием элемента  $v_4^{(l)}$  в  $(k-1)$ -м кадре и не совпадать с одним или двумя окрестными элементами в  $k$ -м кадре, и количество информации в  $v_4^{(l)}$  возрастет.

При смещении объекта элемент  $v_4^{(l)}$  будет иметь различные состояния с элементом  $v_3^{(l)}$  в  $(k-1)$ -м кадре и с одним или двумя окрестными элементами в  $k$ -м кадре,

и количество информации в элементе  $v_4^{(l)}$  еще более увеличится. Причем, чем больше корреляция между соседними кадрами, что характерно для видеопоследовательностей, полученных неподвижным датчиком, тем больше информации будет в элементе  $v_4^{(l)}$  при смещении объектов.

Сравнивая значения вычисленной величины количества информации в элементе изображения с порогом, определяем принадлежность данной точки контуру. Значение порога  $h$  определяют по вычисленной величине количества информации в элементе  $v_4^{(l)}$   $l$ -го РДИ:

$$h = \frac{I(v_4^{(l)} = M_i^{(l)} | v_1^{(l)} = M_i^{(l)}, v_2^{(l)} = M_j^{(l)}, v_3^{(l)} = M_j^{(l)})}{2} + \frac{I(v_4^{(l)} = M_i^{(l)} | v_1^{(l)} = M_j^{(l)}, v_2^{(l)} = M_j^{(l)}, v_3^{(l)} = M_i^{(l)})}{2} \quad (6)$$

Предложенный метод определения контуров движущихся объектов требует небольших вычислительных ресурсов, определяемых операциями сравнения с тремя окрестными элементами.

Результатом такой обработки изображений является совокупность контурных точек, в которых происходит движение одного или нескольких объектов интереса, а также точек, представляющих собой шум, вызванный, например, движением листьев на деревьях, бликами и т.п. Следует отметить, что контурные точки, принадлежащие объекту интереса, являются связными и имеют большую протяженность, чем контурные точки, представляющие собой шум.

При решении задачи выделения объекта интереса необходимо выделить точки, которые наилучшим образом представляют объект. Для этого изображение разбивается на блоки и одновременно с определением контурных точек выполняется анализ указанных блоков на наличие в них связанных пикселей, составляющих часть предполагаемого контура движущего объекта. Размер блока выбирается в соответствии с данными о максимальной скорости движения и минимальной площади

объекта. Решение о присутствии объекта интереса на изображении принимается по наличию связанных пикселей в блоках.

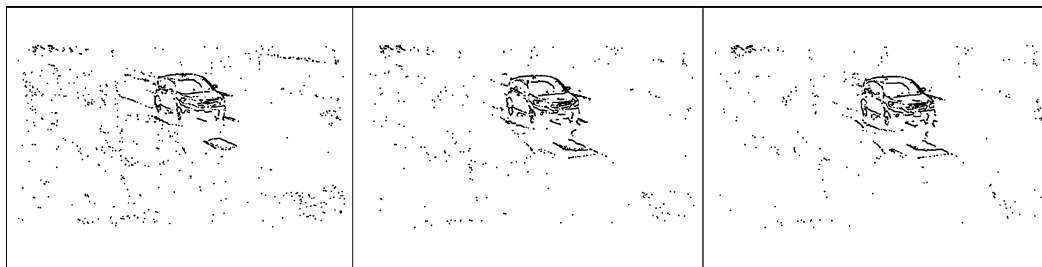
Координаты движущегося объекта соответствуют центру прямоугольного окна, в которое вписывается объект интереса. Последовательность координат движущихся объектов будет представлять собой траекторию движения. При выполнении исследования эффективности разработанного метода использовались несколько видеопоследовательностей средней длительностью 300 кадров.

Для оценки качества выделения движущегося объекта вычислялась вероятность правильного выделения объекта интереса и среднеквадратическая ошибка (СКО) определения координат объекта интереса. На рис. 3 показан фрагмент видеопоследовательности, в котором три соседних кадра совмещены для визуальной оценки передвижения объекта. Ниже на рис. 4 приводятся контурные (а) и выделенные прямоугольным окном объекты (б).



Рис. 3. Совмещенные кадры видеопоследовательности

В результате выделения объектов интереса возникают ошибки двух типов: на контурном изображении объект интереса определен (как совокупность точек шума), в то время как на реальном изображении он отсутствует; на контурном изображении объект не выделен (например, в результате частичного заслонения другим объектом), а на реальном изображении он имеется.



а) Контурные изображения



б) Изображения с выделенными объектами

Рис. 4. Кадры видеопоследовательности с выделенными объектами

Вероятность правильного выделения объекта интереса определялась как отношение количества верно выделенных объектов ( $N_B$ ) к общему количеству движущихся объектов ( $N_O$ ) в видеопоследовательности:

$$P_{np} = \frac{N_B}{N_O} \quad (7)$$

Для приведенной видеопоследовательности общая вероятность правильного определения объекта интереса составила 0,75, что вызвано скрытием объекта интереса препятствием. Размеры объектов интереса варьировались от 30x30 до 120x120 пикселей. Количество одновременно присутствующих объектов в кадре менялось на протяжении всей видеопоследовательности, но не превышало двух. Зависимость вероятности правильного выделения объекта в тестовой видеопоследовательности от номера кадра представлена на рис. 5.

Кадры, в которых вероятность правильного выделения объекта интереса равна нулю, соответствуют ситуации с заслонением объекта интереса. Для других видеопоследовательностей, в которых не происходит скрытие объектов интереса, величина правильного выделения достигает 0,91±0,95.

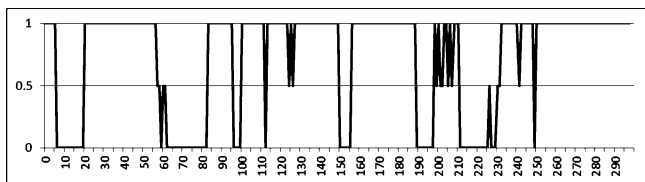


Рис. 5. Зависимость вероятности правильного выделения объекта от номера кадра

Также для оценки качества работы алгоритма вычислялась нормированная среднеквадратическая ошибка (СКО) оценки координат движущегося объекта:

$$СКО = \frac{1}{L} \sqrt{\sigma_x^2 + \sigma_y^2}, \quad (8)$$

где  $\sigma_x$ ,  $\sigma_y$  — смещение координат центра движущегося объекта по горизонтали и вертикали от их истинного значения,  $L$  — половина диагонали прямоугольного окна, обрамляющего объект интереса в идеальном случае.

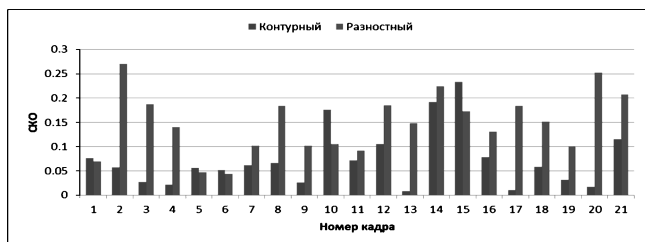


Рис. 6. Сравнение СКО определения координат движущихся объектов разными методами

При выполнении исследования сравнивались результаты СКО оценки координат разработанным и известным разностным методом [1, 3]. Результаты сравнения СКО двумя методами на нескольких кадрах последовательности (рис. 3) представлены на рис. 6. При оценке всей видеопоследовательности из 300 кадров СКО разработанного контурного метода меньше разностного в 2 раза. Для большинства видеопоследователь-

ностей выигрыш в СКО для разработанного метода по сравнению с разностным составляет 1,5÷2,5.

Из приведенных результатов следует, что предложенный метод выделения движущихся объектов в видеопотоке, основанный на вычислении количества информации в элементах изображения, позволяет определить координаты движущегося объекта с точностью, превышающей разностный метод. Разработанный метод требует небольших вычислительных ресурсов, что делает возможным его применение при обработке данных в реальном масштабе времени.

## Литература

1. Яне Б. Цифровая обработка изображений. — М.: Техносфера, 2007. — 584 с.
2. Методы автоматического обнаружения и сопровождения объектов. Обработка изображений и управление / Б.А. Алпатов, П.В. Бабаян, О.Е. Балашов, А.И. Степашкин. — М.: Радиотехника, 2008.-176 с.
3. Богословский А.В. Обработка многомерных сигналов. Линейная многомерная дискретная обработка сигналов. Методы анализа и синтеза. / А.В. Богословский, Е.А. Богословский, И.В. Жигулина, В.А. Яковлев. — М.: Радиотехника, 2013. — 168 с.
4. Трифонов А.П. Обработка динамических изображений. Обнаружение и оценка параметров движения. / А.П. Трифонов, Р.В. Куцов — LAP LAMBER Academic Publishing, Germany, 2011. — 156 с.
5. Петров Е.П. Метод синтеза математических моделей видеоизображений на основе многомерных цепей Маркова / Е.П. Петров, Е. В. Медведева, А.П. Метелев // Нелинейный мир, № 4, 2011. — С.213-231.
6. Карлушин К.А. Метод выделения движущихся объектов в видеоизображениях на основе трехмерных цепей Маркова / К.А. Карлушин, Е. В. Медведева // Т-Comm. Телекоммуникации и транспорт. № 9, 2013. — С.94-97.
7. Карлушин К.А. Метод выделения движущихся объектов на основе пространственно-временной модели изображения / К.А. Карлушин, Е.В. Медведева // Труды 69-й межд. конф., посвященной Дню радио «Радиоэлектронные устройства и системы для инфокоммуникационных технологий»: — М.: 2014 — С.136-140.

## A METHOD OF MOVING OBJECTS DETECTION IN VIDEO STREAM AND ESTIMATION ACCURACY OF OBJECTS COORDINATES DETERMINATION

*Medvedeva E. V., Karlushin K. A.*

A method of moving objects detection in video stream is offered. The method is based on the representation of a sequence of images by three-dimensional Markov chain and moving objects contours extraction by calculating the magnitude of information amount in the image elements. The accuracy of moving objects coordinates determination is estimated. The developed method requires low computational resources, which makes possible its use in the data processing in real time.

УДК 621.391.268

## ПОВЫШЕНИЕ ЭФФЕКТИВНОСТИ СЕЛЕКЦИИ ОБЪЕКТОВ В СИСТЕМАХ ВИДЕОСЛЕЖЕНИЯ ПРИ НАЛИЧИИ ЯРКОСТНЫХ ПОМЕХ

*Муравьев В.С., к.т.н., доцент кафедры автоматики и информационных технологий в управлении  
Рязанского государственного радиотехнического университета, e-mail: aitu@rsreu.ru;*

*Фельдман А.Б. к.т.н., н.с. кафедры автоматики и информационных технологий в управлении  
Рязанского государственного радиотехнического университета.*

**Ключевые слова:** объект, яркостная помеха, обнаружение, видеослежение, надежность, алгоритм селекции, система признаков.

### Введение

В настоящее время технологии слежения за объектами в последовательности видеок кадров находят широкое применение в системах мониторинга транспорта, в бортовых системах летательных аппаратов, в сфере безопасности и других приложениях. Основными задачами, решаемыми системами видеослежения, являются обнаружение, измерение координат и прогнозирование положения объекта интереса. К алгоритмическому обеспечению систем видеослежения зачастую предъявляется требование устойчивой работы при наблюдении объектов в присутствии различного рода помех.

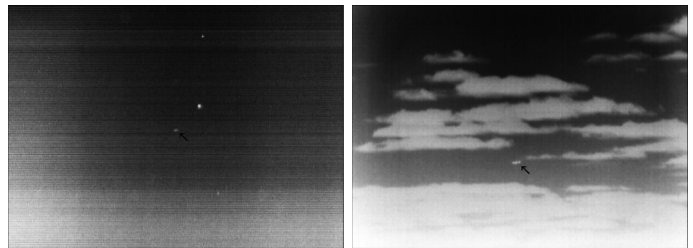
В данной работе в качестве объектов интереса выступают летательные аппараты, наблюдаемые на облачном фоне. Предполагается, что в кадр могут попадать хорошо различимые яркостные образования, сходные с изображением объекта, что может приводить к его потере и срыву процесса слежения. В качестве подобных помех могут выступать изображения птиц и локальные неоднородности фона. Особенно остро данная проблема стоит при наблюдении малоконтрастных удаленных объектов в инфракрасном диапазоне. Система видеослежения может устанавливаться как стационарно, так и размещаться на подвижном носителе. Примеры изображений с характерной фоноцелевой обстановкой представлены на рис. 1.

Одним из способов повышения качества слежения является построение набора признаков, который позволит с большей достоверностью по сравнению с существующими решениями отличить объект от помех. Предполагается, что при обработке изображения с помощью какого-либо известного алгоритма обнаружения были выделены бинарные маски объекта и помех, определяющих их конфигурацию и местоположение [1-3]. При этом помехам на бинарном изображении будут соответствовать сегменты, схожие с объектом интереса и мешающими его наблюдению.

На практике в качестве признаков сегментов используются площадь, габаритные размеры, отношение сторон, средняя яркость. Будем считать, что отмеченные признаки составляют базовый набор. На основании только базового набора не всегда представляется воз-

*Рассматривается возможность повышения надежности слежения за объектами по данным видеонаблюдения при наличии яркостных помех путем построения алгоритма селекции на основе предложенной системы признаков. Разработаны количественные критерии эффективности работы системы видеослежения. Представлены результаты сравнительных экспериментальных исследований, демонстрирующие положительный эффект, связанный с введением новых признаков.*

можным однозначно отличить объект от помехи. В работах [4-6] описывается применение для анализа изображений следующих признаков: моменты изображения, эксцентриситет, направление главной оси, округлость, координаты центра тяжести сегмента и ряд других. Выбранный набор признаков, как правило, подвергается нормализации и подается на классификатор, который принимает решение о наличии интересующего объекта на изображении. Однако многие признаки в реальных ситуациях не способны дать релевантное описание характеристик объекта, а применение сложных классификаторов предполагает решение сопутствующих проблем, возникающих при обучении алгоритма, которые не всегда возможно успешно решить на практике.



*Рис. 1. Примеры изображений с характерной фоноцелевой обстановкой.  
Объект интереса отмечен стрелкой*

Поэтому разработка надежной системы признаков, позволяющих более эффективно селективировать объекты интереса при сохранении простой структуры классификатора, представляется весьма актуальной проблемой.

### Описание признаков

При построении системы признаков необходимо отталкиваться от того факта, что объекты интереса могут быть удалены от наблюдателя, а формируемое изображение, как правило, не содержит выраженных деталей. Данные особенности задачи не позволяют эффективно использовать простейшие характеристики, описываю-

щие форму и яркость объекта. Будем учитывать, что при сравнительно высокой частоте поступления видеокладов можно пренебречь межкадровым изменением масштаба и ориентации объекта. Таким образом, жесткое требование обеспечения полной инвариантности признаков к действию ракурсных геометрических преобразований может быть исключено из рассмотрения.

Вместе с тем анализ информации необходимо производить в темпе поступления данных, что накладывает ограничения на вычислительную сложность применяемых методов и приводит к необходимости поиска компромисса между качеством описания изображения объекта и вычислительной трудоемкостью его получения.

В ходе исследований были предложены пять дополнительных признаков, которые целесообразно использовать совместно с базовым набором. Первые два признака являются яркостными и вычисляются на основе исходного изображения, а оставшиеся строятся на основе бинарных масок и их можно отнести к морфологическим характеристикам сегментов.

Первым признаком является минимум разностной корреляционной функции, отражающей меру сходства эталона и найденных на текущем кадре сегментов. Предполагается, что ранее на этапе выделения и обнаружения получено изображение отслеживаемого объекта  $h(x, y)$  размерностью  $(W_x^{(s)}, W_y^{(s)})$ , рассматриваемое в качестве эталона, координаты которого задаются множеством  $H = \{(x, y) | x = \overline{1, W_x^{(s)}}, y = \overline{1, W_y^{(s)}}\}$ . Отыскивается местоположение и размеры  $(W_x^{(i)}, W_y^{(i)})$ ,  $i = \overline{1, n}$  каждого бинарного сегмента в текущем кадре, где  $n$  – число сегментов в текущем кадре. Тогда, располагая исходным полутоновым изображением для каждого сегмента, можно определить изображение, задаваемого функцией яркости  $l_i(x, y)$ ,  $(x, y) \in H^{(i)}$ ,

$$H^{(i)} = \{(x, y) | x = \overline{1, W_x^{(i)}}, y = \overline{1, W_y^{(i)}}\}.$$

Мера сходства эталона с  $i$ -м сегментом находится путем вычисления минимальных значений разностных корреляционных функций вида (1) в соответствующей зоне поиска  $G^{(i)} = \{(\alpha, \beta) | \alpha = \overline{1, W_x - W_x^{(s)} + 1},$

$$\beta = \overline{1, W_y - W_y^{(s)} + 1}\}$$

$$Km_i = \min_{(\alpha, \beta) \in G^{(i)}} \sum_{(x, y) \in H} |l_i(\alpha + x, \beta + y) - h(x, y)|. \quad (1)$$

В том случае, если размеры эталона превышают размеры сегмента, то признак (1) не вычисляется. Для улучшения качества сопоставления сегментов с эталоном границы изображения каждого сегмента, предлагается расширить на некоторую величину, зависящую от размеров объекта.

Следующий признак строится на основе оценивания распределения энергий градиентов по направлениям. Процедура расчета этого признака во многом схожа с той, что используется в алгоритме описания особенностей изображения SIFT [7] на этапе построения дескрипторов, и включает следующие основные этапы: сглаживание исходного изображения по найденной бинарной

маске сегмента; получение поля градиентов сглаженного изображения; вычисление модуля и направления градиента в каждой точке; оценка распределения энергии градиентов по направлениям; нормирование полученных зависимостей. Предполагается, что данный признак способен дать достаточно полное описание локальных особенностей изображения объекта.

Сглаживание исходного изображения производится гауссовским фильтром с небольшой апертурой размером 5-7 пикселей с целью уменьшения влияния шума. Для каждой точки  $(x, y)$  сглаженного изображения объекта-кандидата численными методами определяется модуль  $m_g(x, y)$  и направление  $\theta_g(x, y)$  градиента. В качестве численной аппроксимации градиента используются конечные разности. Все возможные угловые направления векторов градиента, лежащие на отрезке  $[0; 2\pi)$ , разбиваются на  $N_g$  интервалов (рис. 2). Если обозначить через  $M_g(n_g)$  множество всех точек, в которых направление градиента попадает на интервал разбиения  $n_g = \overline{1, N_g}$ , тогда оценить распределение энергии градиентов по направлениям можно следующим образом

$$G(n_g) = \sum_{(x, y) \in M_g(n_g)} m_g(x, y). \quad (2)$$

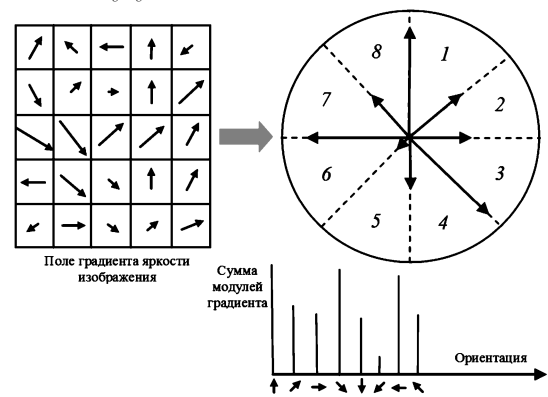


Рис. 2. Оценивание распределения энергии градиентов по направлениям при  $N_g = 8$

Степень сходства оценок распределений, полученных для отслеживаемого объекта  $G_o$  и  $i$ -ого сегмента, находится с помощью коэффициента Бхаттачарья [8], значение которого нормировано и лежит в диапазоне от нуля до единицы:

$$Kb_i = \sum_{n_g} \sqrt{\overline{G_o(n_g)} \overline{G_i(n_g)}}, \quad (3)$$

где  $\overline{G_o(n_g)}, \overline{G_i(n_g)}$  – нормированные оценки распределений. Значение  $N_g$  выбиралось равным от 8, 10 или 12.

Третий признак вычисляется по бинарной маске сегмента, для чего строится описывающий сегмент прямоугольник, который разделяется на восемь треугольников одинаковой площади, общей вершиной которых является центр прямоугольника. В каждой такой области осуществляется подсчет числа попадавших в нее пикселей сегмента. Затем каждый элемент полученного массива делится на площадь сегмента. Таким образом, формируется 8-отсчетная гистограмма, показывающая распре-



деление точек сегмента по областям прямоугольника. Полученный массив фактически отражает меру симметрии бинарной маски. Мерой сходства гистограмм, построенных описанным образом для сегментов и объекта слежения, служит коэффициент Бхаттачарья.

Следующие признаки математически описывают особенности контура бинарного изображения сегмента. В качестве четвертого признака выступает функция перегиба контура. Для получения массива координат граничных пикселей сегмента путем непосредственной обработки бинарной маски, используется алгоритм последовательного обхода контура, описанный в [9]. Для уменьшения влияния шума при описании граничных точек довольно часто применяются подходы к аппроксимации контура, однако для вычисления описываемого признака этого не требуется, что является несомненным достоинством.

Процедура расчета величины перегиба контура  $P(m)$  в точке  $m$  включает нескольких шагов [9]:

- выбор двух точек последовательности  $P(m-L)$  и  $P(m+L)$ , равноудаленных от  $m$  на  $L$  точек;
- определение углов наклона в левую  $\varphi_l$  и правую  $\varphi_r$  сторону от точки  $m$  (рис. 3);
- вычисление  $P(m)$  как разности между углами наклона  $\varphi_l$  и  $\varphi_r$ .

Контур сегмента описывается одномерной функцией перегиба путем вычисления значений  $P(m)$  для всех его точек.

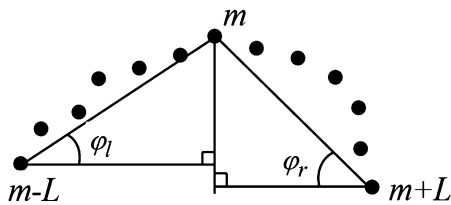


Рис. 3. Вычисление величины перегиба контура при  $L = 6$

Степень схожести функций перегиба для объекта слежения и сегментов, найденных в текущем кадре, предлагается определять, основываясь на значении минимума разностной корреляционной функции. Значение  $L$  на практике выбиралось из диапазона [3; 6].

Последний из дополнительных признаков также рассчитывается по контурным точкам сегментов, но иным способом. Координаты найденных контурных точек преобразуются из декартовой системы координат в полярную. Затем весь диапазон углов разбивается на заданное число интервалов, после чего вычисляется оценка распределения по выражению аналогичному (2), но с тем отличием, что модуль градиента заменяется на радиус-вектор, а фаза – на угловое направление точек контура. Указанный признак служит в качестве дескриптора формы и находит широкое распространение в сис-темах технического зрения [10]. Как и ранее, в роли меры сходства гистограмм выступает коэффициент Бхаттачарья.

#### Алгоритм селекции объектов

После нахождения признаков и мер их сходства возникает необходимость в построении алгоритма селек-

ции, который бы позволил совместно использовать как предлагаемый, так и базовый наборы признаков. Следует отметить, что алгоритм селекции на основе базового набора был разработан ранее и апробирован на практике [11]. Поэтому рассматривалась возможность интеграции нового набора признаков в уже существующий алгоритм, доказавший свою работоспособность. Укажем основные этапы предлагаемого алгоритма.

1) Формирование базового набора из  $M_\delta$  признаков для  $i$ -го найденного сегмента при  $i = \overline{1, n}$ ,  $n$  – число сегментов.

2) Нахождение относительного отклонения  $\delta^{(i)} = [\delta_s^{(i)}]$ ,  $s = \overline{1, M_\delta}$  между сглаженными значениями признаков объекта, рассчитанными на прошлом кадре, и текущими значениями признаков из базового набора.

3) Формирование массивов бинарных параметров  $\mathbf{B}_\delta^{(i)} = [B_{\delta_s}^{(i)}]$ ,  $s = \overline{1, M_\delta}$  путем сравнения  $\delta^{(i)}$  с заранее заданными порогами.

4) Формирование набора из  $M_\delta$  разработанных дополнительных признаков для всех найденных сегментов;

5) Нахождение мер сходства между признаками объекта, найденными на прошлом кадре, и текущими признаками из дополнительного набора.

6) Проверка мер сходства и формирование массива бинарных параметров  $\mathbf{B}_\delta^{(i)} = [B_{\delta_s}^{(i)}]$ ,  $s = \overline{1, M_\delta}$  для  $i$ -го сегмента,  $i = \overline{1, n}$ .

а) Для гистограммных признаков значения из массива  $\mathbf{B}_\delta^{(i)}$  находятся путем непосредственного сравнения полученных значений мер с заданными порогами. Значения порогов выбираются из диапазона [0,85; 0,95].

б) Для признаков, сравниваемых на схожесть с эталоном, соответствующие значения из массива  $\mathbf{B}_\delta^{(i)}$  устанавливаются в единицу при выполнении неравенства:

$$Km_i \leq \min_i \{Km_i\} + (\max_i \{Km_i\} - \min_i \{Km_i\})k_f, \quad (4)$$

$i = \overline{1, n}$

где  $Km_i$  – минимум разностной корреляционной функции для признака  $i$ -го сегмента,  $k_f$  – коэффициент, принимающий значения от 0 до 1. В случае несоблюдения условия (4) бинарный параметр приравнивается нулю.

7) Исключение  $i$ -го сегмента из списка, если хотя бы одно значение бинарного параметра из массивов  $\mathbf{B}_\delta^{(i)}$  и  $\mathbf{B}_\delta^{(i)}$  равно нулю;

8) За объект принимается сегмент, имеющий минимальное евклидово расстояние между своим центром и центром строка слежения (область в которую координаты объекта попадают с большой вероятностью в текущем кадре);

9) Межкадровое сглаживание численных значений базовых признаков с помощью экспоненциального фильтра.

10) Временная фильтрация распределений и функций перегиба, используемых в качестве признаков и смена эталонов. Введение этапа межкадровой фильтрации значений признаков способствует увеличению качества работы алгоритма.

11) При отсутствии объекта в стробе положение области поиска прогнозируется на следующий кадр. Для формирования строба слежения применяются известные подходы, использующие математический аппарат калмановской фильтрации [12].

### Экспериментальные исследования

При проведении сравнительных экспериментальных исследований ставилась задача оценки положительного эффекта, связанного с использованием выработанных признаков в алгоритме селекции при слежении за объектом интереса. В силу большого многообразия возможных сцен, условий наблюдения, характеристик движения объекта интереса, которые в совокупности нельзя учесть в моделях фоноцелевой обстановки, можно заключить, что показатели эффективности алгоритма должны вычисляться с использованием статистики, собранной при обработке достаточно большого числа тестовых видеок кадров. С точки зрения специфики прикладных задач, решаемых с помощью систем видеослежения, важно оценить, во-первых, качество обнаружения и селекции объектов интереса, а, во-вторых, надежность прослеживания объектов в последовательности видеок кадров. Зачастую требуется найти точность измерения координат обнаруженного объекта.

Важным показателем является частота правильной селекции. Данный критерий определяется как отношение числа кадров видеопоследовательности, на которых сегмент был правильно отнесен к объекту к общему числу кадров, где объект присутствовал. Желаемым результатом является сравнительно близкое совпадение эталонных координат центра объекта в текущем кадре  $(x_{об}(n), y_{об}(n))$ , заданных человеком-экспертом, с результатами измерения  $(\tilde{x}_{об}(n), \tilde{y}_{об}(n))$ . Сегмент считается правильно отнесенным к объекту в текущем кадре, если разность между эталонными и измеренными координатами не превышает некоторой заданной величины, связанной с размерами объекта и точностью измерения координат. В качестве такой величины принимается размер диагонали  $d_{об}(n)$  прямоугольника, ограничивающего эталонное изображение объекта в данном кадре видеопоследовательности.

Таким образом, объект считался правильно селективным в кадре  $n$  при выполнении неравенства:

$$\sqrt{(x_{об}(n) - \tilde{x}_{об}(n))^2 + (y_{об}(n) - \tilde{y}_{об}(n))^2} \leq 1 + d_{об}(n). \quad (5)$$

Тогда частота правильной селекции может быть найдена как  $\eta_{сел} = N_{ис} / N_{\max}$ , где  $N_{ис}$  – количество кадров, на которых удовлетворяется условие (5),  $N_{\max}$  – число кадров в видеопоследовательности, содержащих объект.

Точность слежения за объектом можно охарактеризовать при помощи СКО ошибок измерения координат. Оценивать точность измерения координат объекта имеет смысл только в том случае, если сегмент был правильно отнесен к объекту в текущем кадре. Целесообразно ввести в рассмотрение СКО совокупной ошибки измерения координат  $\sigma$  определяемой как среднее геометрическое СКО ошибок измерения по отдельным координатам.

В качестве показателя, характеризующего эффективность прослеживания объекта в последовательности кадров, будем рассматривать частоту срывов слежения. Под срывом слежения будем понимать превышение ошибки измерения координат объекта размеров строба слежения хотя бы по одной координате. При возникновении этого события формировалось новое целеуказание на основе имеющихся эталонных данных, и система видеослежения перезапускалась. Число срывов  $K_{ср}$  определяется количеством кадров, на которых координаты объекта не попадают в строб, но при этом попадают в него на предыдущем кадре. Зная общее число срывов и число кадров в видеопоследовательности  $N_{\max}$ , можно оценить частоту срывов слежения  $\eta_{ср} = K_{ср} / N_{\max}$ . Другим показателем, характеризующим качество прослеживания объекта, является отношение средней длительности периодов слежения к длине видеосюжета  $\eta_{слеж}$ .

Для расширения экспериментальной базы было разработано специализированное программное обеспечение (ПО), позволяющее проводить полунатурное моделирование яркостных помех. При изучении влияния помех на работу алгоритма селекции объектов выбирались сюжеты с подходящим типом фоноцелевой обстановки, на которые затем накладывались яркостные помехи. Используя созданное ПО, имелась возможность указать количество помех в серии, стартовый кадр, период следования помехи.

В качестве исходного видеоматериала для проведения сравнительных исследований служила база натуральных видеосюжетов, которая содержала шесть видеопоследовательностей общей продолжительностью более 8000 кадров. На четырех сюжетах в качестве фоновой составляющей выступало ясное небо с неявно выраженной облачностью, остальные сюжеты содержали облачность с контрастными яркостными перепадами. Один из сюжетов содержал постепенно удаляющийся от наблюдателя объект, очень близкий по форме и размерам к помехе на изображении. Размеры объектов на протяжении времени наблюдения изменялись от 10x10 до 40x40 пикселей, а скорость межкадрового сдвига изображения не превышала 15 пикселей.

Для получения бинарных изображений использовался алгоритм обнаружения объектов, основанный на пространственной фильтрации и адаптивной пороговой обработке, изложенный в [1]. В указанный алгоритм был внесен ряд изменений, связанный с коррекцией порога бинаризации. Так, значение порога сглаживалось во времени и фиксировалось на постоянном уровне при его резком возрастании. Данная модификация позволила улучшить качество обнаружения при наличии в зоне обработки большого числа ярких пикселей.

С целью увеличения объема статистических данных проводилось усреднение показателей эффективности селекции и прослеживания объекта при 20 запусках алгоритма с разными параметрами поведения помехи для каждого тестового сюжета. Моделировалось появление серий из двух и четырех помех с периодом следования в несколько сотен кадров. Показатели вычислялись на осно-

ве обработки каждого видеосюжета и затем усреднялись на всей выборке видеосюжетов с учетом их длительности.

Обобщенные значения показателей эффективности для серий из двух помех сведены в табл. 1. В таблице использованы следующие условные обозначения признаков из дополнительного набора: «1» – сходство с эталонным изображением объекта на основе минимума разностной КФ; «2» – распределение площади сегмента по направлениям; «3» – оценка распределения энергии градиентов полутонового изображения по направлениям; «4» – функция перегиба контура сегмента; «5» – гистограммное описание формы контура (дескриптор формы).

На основе анализа полученных результатов было выявлено, что наилучшие результаты обеспечивались

при использовании признаков «1», «2» и «4». В табл. 1 приведены показатели эффективности при использовании парных комбинаций признаков. Использование большего числа признаков было признано нецелесообразным, так как приводит к заметному росту вычислительной трудоемкости при сохранении, а в ряде случаев и ухудшении значений анализируемых критериев.

В табл. 2 приведены результаты работы алгоритма для серий из четырех помех. Следует отметить, что при одновременном появлении большого числа помех они могут объединяться и накладываться на изображение объекта. Данная ситуация может приводить к увеличению времени нахождения алгоритма в режиме прогноза и устареванию информации об объекте.

Таблица 1. Показатели эффективности слежения для серий из двух помех

Признаки	$\eta_{сст}, \%$	$\sigma$ , пикс.	$\eta_{ср}, \%$	$\eta_{слеж}, \%$
Базовые	93,19	3,1	0,192	35,8
Базовые и дополнительные				
1	97,18	2,1	0,126	42
2	96,31	2,3	0,143	40,6
3	92,89	3,1	0,198	35,4
4	97,02	2,1	0,132	41,7
5	94,95	2,6	0,171	38,5
1, 4	97,9	1,8	0,107	43,1
1, 2	97,43	1,9	0,119	42,4
2, 4	97,64	1,9	0,114	42,7
1, 5	95,39	2,5	0,163	39,2
2, 5	95,12	2,6	0,168	38,8

Таблица 2. Показатели эффективности слежения для серий из четырех помех

Признаки	$\eta_{сст}, \%$	$\sigma$ , пикс.	$\eta_{ср}, \%$	$\eta_{слеж}, \%$
Базовые	90,98	4,4	0,458	19,0
Базовые и дополнительные				
1	94,54	3,3	0,382	22
2	93,82	3,6	0,393	20,4
3	89,45	4,6	0,462	17,7
4	94,04	3,5	0,397	21,3
5	91,27	4,2	0,419	19,6
1, 2	94,19	3,4	0,376	21,5
1, 4	95,14	3,1	0,357	22,9
2, 4	94,93	3,2	0,363	22,4

## Заключение

При проведении исследований анализировались показатели эффективности работы системы видеослежения с использованием предложенного алгоритма селекции объектов и дополнительной системы признаков. Ситуация усложнялась наличием на исходном изображении шума, неравномерного фона и различным числом яркостных помех. Форма наблюдаемых на изображениях объектов зачастую была схожа с формой яркостных образований, порожденных помехами, а сами изображения зашумлены. В этих условиях интегральные показатели, к которым можно отнести гистограммные характеристики, оказались менее информативными. Признаки, обеспечивающие лучшие результаты, обладали большей помехоустойчивостью и опирались на использование всей доступной информации об

изображении объекта и особенностях его контура.

На основе полученных данных были сделаны следующие выводы:

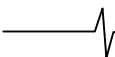
– эффективность слежения уменьшается с ростом числа помех, что связано с увеличением продолжительности заслонения объекта помехой и устареванием информации об объекте;

– наибольшей эффективностью обладают признаки на основе функций перегиба контура и сходства с эталонным изображением объекта;

– точность измерения координат объекта возрастает на 20-40%;

– частота срывов слежения уменьшается в 1,5-2 раза.

В целом алгоритм селекции, совместно использующий базовый и разработанный наборы признаков, продемонстрировал достаточно высокие характеристики,



нетребователен к вычислительным ресурсам и может применяться на практике в составе систем видеослежения реального времени.

Работа выполнена при финансовой поддержке гранта Президента РФ для молодых российских ученых-кандидатов наук МК-958.2013.10.

#### Литература

1. Алпатов Б.А., Блохин А.Н., Муравьев В.С. Алгоритм обработки изображений для систем автоматического сопровождения воздушных объектов // Цифровая обработка сигналов. – 2010. – №4. – С.12-17.

2. Бабаян П.В., Фельдман А.Б. Определение местоположения объекта на изображении в системах технического зрения мобильных робототехнических комп-лексов // Вестник Рязанского государственного радиотехнического университета. – 2011. – № 38. – С. 19-25.

3. Алпатов Б.А., Бабаян П.В., Смирнов С.А. Автоматическое слежение за объектами при отсутствии априорных сведений о фоноцелевой обстановке // Цифровая обработка сигналов. – 2009. – № 3.– С. 52-56.

4. Labonte G., Deck W.C. Infrared target-flare discrimination using a ZISC hardware neural network // Journal of real-time image processing. – Vol. 5, Issue 1. – 2010. – PP. 11-32.

5. Viau C.R. Expendable Countermeasure effectiveness against imaging infrared guided threats // Second International Conference on Electronic Warfare, 2012, pp. 16.

6. Gray G.J., Aouf N., Richardson M.A., Butters B, Walmsley R. An intelligent tracking algorithm for an imaging infrared anti-ship missile // Technologies for optical countermeasures IX. Proc of SPIE. – 2012. – Vol. 8543, 85430L.

7. Lowe D.G. Distinctive image features from scale-invariant keypoints // International journal of computer vision. – 2004. – Vol. 60(2). – PP. 91-110.

8. Reyes-Aldasoro C.C., Bhalerao A. The Bhattacharyya space for feature selection and its application to texture segmentation // Pattern Recognition. – 2006. – Vol. 39, Issue 5. – PP. 812-826.

9. Визильтер Ю.В., Желтов С.Ю., Бондаренко А.В. и др. Обработка и анализ изображений в задачах машинного зрения. – М., 2010. – 672с.

10. Шапиро Л., Стокман Дж. Компьютерное зрение. – М., 2006. – 752с.

11. Способ обработки последовательности изображений для обнаружения и слежения за воздушными объектами: пат. №2419150 РФ / Б.А. Алпатов, П.В. Бабаян, Л.Н. Костяшкин, С.И. Муравьев, В.С. Муравьев, Ю.Н. Романов, В.Н. Эгель; заявитель и патентообладатель ОАО «Государственный Рязанский приборный завод»; заявлен 10.03.2010, опубл. 20.05.2011, Бюл.№14.

12. Bar-Shalom Y., Li X.-R. Kirubarajan T. Estimation with applications to tracking and navigation: theory, algorithms and software, New-York: Wiley, 2001. – 581p.

### IMPROVING THE EFFICIENCY OF OBJECT SELECTION IN VIDEO TRACKING SYSTEMS IN THE PRESENCE OF CLUTTER

*Muraviev V.S.*

In this paper we consider the possibility of increasing the efficiency of video tracking in the presence of clutter by constructing the new object selection algorithm based on the proposed features. The performance measures used in the research of the effectiveness of object tracking are introduced. Results of the comparative experimental studies demonstrated the advantage of the developed algorithm based on the developed set of object features.

### *У в а ж а е м ы е а в т о р ы !*

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

#### **Требования к текстовым материалам и сопроводительным документам:**

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

#### **Требования к иллюстрациям:**

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

УДК 621.397; 004.92

## АЛГОРИТМ УПРОЩЕННОЙ КОМПЕНСАЦИИ ДИСТОРСИИ ПРИ ПРОЕКЦИРОВАНИИ ВИДЕОИЗОБРАЖЕНИЯ НА АСФЕРИЧЕСКИЕ ОТРАЖАЮЩИЕ ПОВЕРХНОСТИ АПРИОРНО НЕИЗВЕСТНОЙ ФОРМЫ

*Холопов И.С., к.т.н., доцент кафедры радиотехнических систем Рязанского государственного радиотехнического университета, e-mail: kholopov.i.s@rsreu.ru.*

**Ключевые слова:** проекционные системы индикации, дисторсия, предискажения, билинейная интерполяция.

### Введение

Проекционные системы индикации (ПСИ) представляют собой комплексные оптические системы, позволяющие отображать в поле зрения оператора дополнительную информацию [1–4], и используются как в военной авиации [1, 2], так и в коммерческих приложениях «виртуальной реальности» [3]. ПСИ просветного типа, в которых видеоинформация накладывается на изображение внешнего пространства, в зарубежной литературе обозначают аббревиатурой HUD (от англ. «Head-Up-Display») [4]. Большинство выпускаемых и перспективных HUD основано на использовании селективного оптического покрытия, наносимого на поверхность оптического светоделительного элемента – комбинера (рис. 1) [1, 4, 5]. На рис. 1:

- 1) зрачковая зона – область расположения глаз наблюдателя;
- 2) комбинер;
- 3) плоскость промежуточного изображения;
- 4) проекционный объектив;
- 5) корректирующий клин;
- 6) генератор двумерных изображений – LCD-матрица (микродисплей).

Для получения минимальных габаритных размеров для ПСИ светоделительная (отражающая) поверхность комбинера должна иметь не плоскую, а вогнутую форму: сферическую или, что более предпочтительно, асферическую [5, 6].

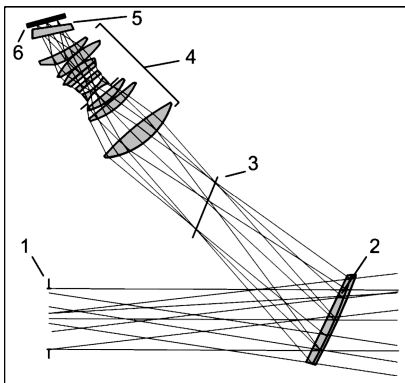


Рис. 1. Оптическая система HUD коллиматорного типа с трассировкой хода световых лучей

*Приведены аналитические выражения для упрощенного формирования предискажений кадра видеоизображения, проецируемого на асферическую отражающую поверхность априорно неизвестной формы.*

Поскольку поверхность комбинера развернута как по отношению к глазам наблюдателя, так и по отношению к проекционному объективу [4–6], то при относительно больших полях зрения, наблюдаемых двумя глазами, и большом выносе выходного зрачка (области размещения глаз наблюдателя), она вносит трансформацию световых пучков и строит их наклонное изображение с дисторсионными искажениями (ДИ), а также аберрациями (отклонениями от гомоцентричности) внеосевых световых пучков – комой и астигматизмом [6], вызывающими размытие и нерезкость изображения.

Практика разработки и расчетов HUD для ПСИ показывает [2, 5, 6], что осуществить компенсацию ДИ применением только оптических элементов не удастся, поэтому необходимо внесение предискажений (ПИ) в изображение, формируемое матрицей микродисплея. Аналитическое описание ДИ, вносимых проекционной системой HUD, затруднено, во-первых, нелинейными функциями описания асферических поверхностей (АП) [7]:

$$Z(S) = \frac{cS^2}{1 + \sqrt{1 - (K+1)c^2S^2}} + A_1S^4 + A_2S^6 + A_3S^8 + A_4S^{10},$$

где  $S = \sqrt{x^2 + y^2}$ ,  $K$  – параметр, определяющий тип поверхности (табл. 1),  $c$  и  $A_1$ – $A_4$  – некоторые постоянные, а, во-вторых, изменяемым положением микродисплеев относительно оптических центров АП комбинера (например, в бинокулярных HUD расстояние между центрами микродисплеев регулируется в соответствии с базой глаз оператора [1–3, 5]).

Таблица 1

$K$	Поверхность
$< -1$	гиперболоид
$-1$	параболоид
$-1 < K < 0$	эллипсоид вращения вокруг большей оси
$0$	сфера
$> 0$	эллипсоид вращения вокруг меньшей оси

В связи с этим, как отмечено в [6], необходимо аналитически описать ДИ, вносимые ПСИ, и внести ПИ таким образом, чтобы окончательная величина дисторсии в пространстве изображений соответствовала принятым

для HUD допускам. Это может быть выполнено, например, путем вычисления коэффициентов полиномиального разложения функций ДИ и ПИ.

**Цель работы:** разработка алгоритма упрощенного формирования ПИ для компенсации ДИ при априорно неизвестной форме АП, на которую проецируется видеоизображение.

### Алгоритмы полиномиальной компенсации ДИ

ДИ оптико-электронных устройств принято описывать комбинацией радиальной и тангенциальной дисторсий [7-9], для компенсации которых на практике используют формулы:

$$\begin{aligned} x_0 &= x'(1 + k_1 r^2 + k_2 r^4) + 2p_1 x' y' + p_2 (r^2 + 2x'^2) + x_c, \\ y_0 &= y'(1 + k_1 r^2 + k_2 r^4) + p_1 (r^2 + 2y'^2) + 2p_2 x' y' + y_c, \end{aligned} \quad (1)$$

где  $r^2 = x'^2 + y'^2$ ,  $x' = x_d - x_c$ ,  $y' = y_d - y_c$ ,  $k_1$  и  $k_2$  – коэффициенты компенсации радиальной дисторсии,  $p_1$  и  $p_2$  – коэффициенты компенсации тангенциальной дисторсии,  $(x_d, y_d)$  – координаты пикселя изображения с ДИ,  $(x_c, y_c)$  – координаты центра дисторсии.

Другим подходом является вычисление поправок  $(\Delta x_0, \Delta y_0)$  к координатам каждого пикселя с использованием полиномов вида [10]:

$$\begin{aligned} x_0 &= x_d + \Delta x_0, \\ y_0 &= y_d + \Delta y_0, \\ Dx_0 &= a_0 + a_1 x' + a_2 y' + a_3 x'^2 + a_4 x' y' + \\ &+ a_5 y'^2 + a_6 x'^3 + a_7 x'^2 y' + a_8 x' y'^2 + a_9 y'^3, \\ Dy_0 &= b_0 + b_1 y' + b_2 x' + b_3 y'^2 + b_4 y' x' + \\ &+ b_5 x'^2 + b_6 y'^3 + b_7 y'^2 x' + b_8 y' x'^2 + b_9 x'^3. \end{aligned} \quad (2)$$

Для определения параметров компенсации дисторсии  $k_1, k_2, p_1, p_2, x_c, y_c$  и коэффициентов М.М. Русинова [10, 11]  $a_i, b_i, i = \overline{0, 9}$ , в выражениях (1) и (2) соответственно используется эталонное тестовое изображение (например, приводимое в [9] или [11]) и изображение, полученное с диагностической камеры, установленной в зрачковой зоне 1 (рис. 1). Вычисление параметров основано на составлении переопределенной системы из  $M$  пар уравнений вида

$$\begin{aligned} \begin{pmatrix} x_{0i} \\ y_{0i} \end{pmatrix} &= F \left( \begin{pmatrix} x_{di} \\ y_{di} \end{pmatrix} \right), \quad i = \overline{1, M}, \end{aligned} \quad (3)$$

из выражений (1) или (2) и ее решении численными методами Ньютона или Левенберга-Маквардта [8, 11]. В рассматриваемых формулах (1) и (2) число неизвестных параметров, определяющих ДИ, равно 6 и 20, поэтому для их нахождения требуется решение системы уравнений (3), составленной соответственно из  $M > 3$  и  $M > 10$  пар уравнений.

Недостатком методов (1) и (2) является сложность коррекции ДИ во внелабораторных условиях (при отсутствии диагностической камеры и стенда измерения параметров ДИ). Поскольку значения параметров компенсации дисторсии  $k_1, k_2, p_1, p_2, x_c, y_c$  в (1) и  $a_i, b_i, i = \overline{0, 9}$ , в (2), как показано в [5, 8, 11], отличаются друг от друга на порядки, это затрудняет подбор их значений методом перебора с фиксированным шагом.

Также на практике для формирования ПИ удобно знать аналитическое выражение не для  $F(\cdot)$ , а для обратного функционала  $F^{-1}$ :

$$\begin{pmatrix} x \\ y \end{pmatrix} = F^{-1} \left( \begin{pmatrix} i \\ j \end{pmatrix} \right),$$

чтобы задавая целочисленные значения координат пикселей  $(i, j)$  кадра видеоизображения с ПИ вычислять соответствующие им дробные координаты  $(x, y)$  пикселей в исходном кадре. Для нелинейных выражений (1) и (2) нельзя аналитически задать обратный функционал  $F^{-1}(\cdot)$ , реализующий однозначное отображение  $(i, j)$  в  $(x, y)$ .

### Алгоритм упрощенной компенсации ДИ

Идея упрощенного формирования ПИ основана на характерных для HUD типах ДИ [2, 6]. Схематично исходное полутоновое изображение показано на рис. 2, а; ДИ, вносимые комбинером – на рис. 2, б; изображение с ПИ – на рис. 2, в, где  $W$  и  $H$  – соответственно ширина и высота кадра видеоизображения в пикселях.

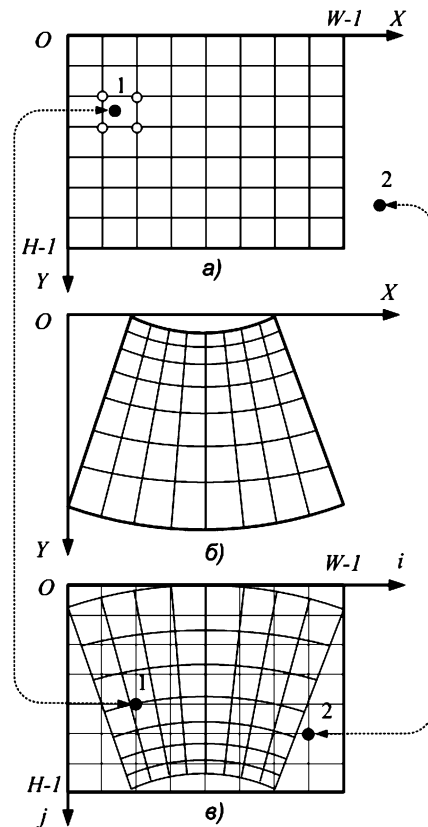


Рис. 2. Кадры видеоизображения: а) исходный кадр; б) отображение кадра на АП; в) кадр с ПИ

На рис. 2 можно выделить 3 типа характерных для HUD ДИ:

1) искажения типа «парабола» в направлении оси  $OY$ ; данный тип ДИ при  $p_1 = p_2 = 0$  может быть описан уравнениями (1), если для координат центра радиальной дисторсии справедливы условия:

$$\begin{cases} x_c = W / 2, \\ y_c > H; \end{cases}$$

2) искажения типа «трапеция» в направлении оси  $OX$ ;

3) неравномерность масштаба в направлении оси  $OY$ .

ДИ типа 2) являются результатом проективного преобразования координат, а ДИ типов 1) и 3) – полиномиального [12].

Следовательно, аналитической функции, описывающей ДИ, необходимо сформировать ПИ таким образом, чтобы изображение в поле зрения оператора имело прямоугольную форму с равномерным масштабом по оси  $OY$ .

Алгоритм упрощенной компенсации ДИ заключается в последовательном формировании ПИ, компенсирующих ДИ типов 1), 2) и 3) исходного изображения.

Для формирования ПИ по оси  $OY$  типа «парабола» выполняется преобразование:

$$x_i = i, \quad y_j = \begin{cases} \frac{j - \frac{4b_y H(i-W/2)^2}{W^2}}{1-b_y}, & b_y > 0, \\ \frac{j - b_y H \left[ \frac{4(i-W/2)^2}{W^2} - 1 \right]}{1+b_y}, & b_y < 0, \end{cases} \quad (4)$$

где  $(i, j)$  – целочисленные координаты пикселя изображения с предыскажениями, расположенного на пересечении  $i$ -го столбца и  $j$ -й строки;  $(x_i, y_j)$  – в общем случае дробные координаты пикселя в исходном изображении, соответствующие точке  $(i, j)$  после выполнения (4),  $i = 0, 1, \dots, W-1, j = 0, 1, \dots, H-1$ ;  $b_y$  – коэффициент, определяющий относительную кривизну в направлении оси  $OY$ ,  $b_y \in (-1, 1)$ .

При априорно неизвестной форме АП и,

Для формирования ПИ по оси  $OX$  типа «трапеция» выполняется преобразование:

$$u_i = \begin{cases} x_i / [1 + tr_x (y_{bi} - H) / 2H] + W / 2, & tr_x > 0, \\ x_i / (1 + tr_x y_{bi} / 2H) + W / 2, & tr_x < 0, \end{cases} \quad (5)$$

$$v_j = y_j,$$

где  $tr_x$  – коэффициент, определяющий в зависимости от его знака относительное сжатие верхней или нижней стороны изображения вдоль оси  $OX$ ,  $tr_x \in [-1, 1]$ ,  $(u_i, v_j)$  – координаты пикселя в исходном изображении, соответствующие точке  $(i, j)$  после выполнения (4) и (5).

Для компенсации неравномерности масштаба по оси  $OY$  используется выражение:

$$x_{0i} = u_i, \quad y_{0j} = v_j \left[ 2 - sc_y + 2(sc_y - 1) \cdot j / H \right] / sc_y, \quad (6)$$

где  $sc_y$  – коэффициент, показывающий относительное изменение масштаба вдоль оси  $OY$ ,  $sc_y \in (0, 2)$ .

Алгоритм упрощенного формирования ПИ изображений, проецируемых на АП, заключается в последовательном выполнении преобразований (4)-(6) и нахождения координат  $(x_{0i}, y_{0j})$  в исходном изображении, соответствующих точке  $(i, j)$ . Если для точки  $(x_{0i}, y_{0j})$  одновременно выполняются неравенства  $0 \leq x_{0i} \leq W-1$  и  $0 \leq y_{0i} \leq H-1$  (точка 1 на рис. 2, а и в), то для нахождения яркости пикселя  $I(i, j)$  применяется [12, 13] либо правило «ближайшего соседа», либо билинейная интерполяция по соседним точкам  $I(\lfloor x_{0i} \rfloor, \lfloor y_{0j} \rfloor)$ ,  $I(\lfloor x_{0i} \rfloor + 1, \lfloor y_{0j} \rfloor)$ ,  $I(\lfloor x_{0i} \rfloor, \lfloor y_{0j} \rfloor + 1)$  и  $I(\lfloor x_{0i} \rfloor + 1, \lfloor y_{0j} \rfloor + 1)$ , где оператор  $\lfloor \cdot \rfloor$  означает округление до ближайшего

меньшего целого. В противном случае принимается  $I(i, j) = 0$  (точка 2 на рис. 2, а и в).

При использовании правила «ближайшего соседа» яркость пикселя  $I(i, j)$  находится по формуле:

$$I(i, j) = \begin{cases} I(\lfloor x_{0i} \rfloor, \lfloor y_{0j} \rfloor), & (x_{0i} - \lfloor x_{0i} \rfloor < 0,5) \& (y_{0j} - \lfloor y_{0j} \rfloor < 0,5), \\ I(\lfloor x_{0i} \rfloor + 1, \lfloor y_{0j} \rfloor), & (x_{0i} - \lfloor x_{0i} \rfloor \geq 0,5) \& (y_{0j} - \lfloor y_{0j} \rfloor < 0,5), \\ I(\lfloor x_{0i} \rfloor, \lfloor y_{0j} \rfloor + 1), & (x_{0i} - \lfloor x_{0i} \rfloor < 0,5) \& (y_{0j} - \lfloor y_{0j} \rfloor \geq 0,5), \\ I(\lfloor x_{0i} \rfloor + 1, \lfloor y_{0j} \rfloor + 1), & (x_{0i} - \lfloor x_{0i} \rfloor \geq 0,5) \& (y_{0j} - \lfloor y_{0j} \rfloor \geq 0,5). \end{cases}$$

При использовании билинейной интерполяции яркость пикселя  $I(i, j)$  находится по формуле:

$$I(i, j) = \lfloor I_j + (I_{j+1} - I_j)(y_{0j} - \lfloor y_{0j} \rfloor) \rfloor,$$

где

$$I_j = I(\lfloor x_{0i} \rfloor, \lfloor y_{0j} \rfloor) + (I(\lfloor x_{0i} \rfloor + 1, \lfloor y_{0j} \rfloor) - I(\lfloor x_{0i} \rfloor, \lfloor y_{0j} \rfloor))(x_{0i} - \lfloor x_{0i} \rfloor),$$

$$I_{j+1} = I(\lfloor x_{0i} \rfloor, \lfloor y_{0j} \rfloor + 1) + (I(\lfloor x_{0i} \rfloor + 1, \lfloor y_{0j} \rfloor + 1) - I(\lfloor x_{0i} \rfloor, \lfloor y_{0j} \rfloor + 1))(x_{0i} - \lfloor x_{0i} \rfloor).$$

Для компенсации ДИ, схематично показанных на рис. 2, б, можно ограничиться формулами (4)-(6).

Процедура подбора значений параметров  $tr_x$ ,  $b_y$  и  $sc_y$  является эмпирической и заключается в поочередном (сначала подбирается значение  $tr_x$ , затем –  $b_y$ , затем –  $sc_y$ ) переборе данных параметров с фиксированным шагом 0,005 или 0,01, начиная со значения 1: для компенсации ДИ, приведенных на рис. 2, б,  $tr_x$  декрементируется, а  $b_y$  и  $sc_y$  – инкрементируются. Перебор прекращается либо при достижении максимального визуального комфорта оператора (линии сетки тестового изображения на АП воспринимаются прямыми, параллельными и эквидистантными), либо при максимальной близости тестового изображения и изображения диагностической камеры (мера близости минимизируется по методу наименьших квадратов [12, 13]). В последнем случае при подборе параметра  $tr_x$  минимизируется квадрат ошибки относительно вертикальных линий тестового изображения, при подборе  $b_y$  – относительно горизонтальных линий и при подборе  $sc_y$  – относительно как вертикальных, так и горизонтальных линий.

Недостатками всех рассмотренных алгоритмов формирования ПИ для компенсации ДИ (1), (2) и (4)-(6) являются потеря части полезной информации из кадра видеоизображения (при невыполнении условий  $0 \leq x_{0i} \leq W-1$  и  $0 \leq y_{0i} \leq H-1$ ) и снижение разрешающей способности.

#### Оценка вычислительных затрат

Количество элементарных вычислительных операций (ЭВО) на реализацию преобразований (4)-(6) и определение яркости пикселей в изображении с ПИ размером  $W \times H$  пикселей приведено в табл. 2.

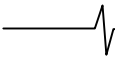


Таблица 2

Операция	«+»	«*»	«/»	ЭВО
Алгоритм «ближайшего соседа»	9WH	9WH	WH	19WH
Алгоритм билинейной интерполяции	18WH	12WH	WH	31WH

В табл. 3 приведена скорость обработки (кадров/с) видеопоследовательности с размером кадра  $W \times H$  пикселей в реальном масштабе времени с использованием разных ПЭВМ и алгоритма билинейной интерполяции.

Таблица 3

Размер кадра, пикс. x пикс.	800x600	1024x768	1280x1024
Pentium Dual CPU E2200, 2,2 ГГц, ОЗУ 1 Гб	14,5	8,7	5,3
Intel Core i5, 2,4 ГГц, ОЗУ 4 Гб	45,9	27,5	16,8

Из табл. 3 видно, что для формирования ПИ в реальном масштабе времени рациональнее вычислять координаты  $(x_{0i}, y_{0i})$  только один раз и затем хранить их в памяти ПЗУ: например, в ячейку с адресом  $2(jW + i)$  записывать координату  $x_{0i}$ , вычисленную в соответствии с (4)-(6), а в ячейку с адресом  $2(jW + i) + 1$  – координату  $y_{0i}$ . Исходя из предположения, что для хранения координат используется ПЗУ с ячейкой памяти 16 бит, минимальный объем ПЗУ, требуемый для разных размеров кадра, приведен в табл. 4.

Таблица 4

Размер кадра, пикс.	800x600	1024x768	1280x1024
Объем ПЗУ, Мбит	7,68	12,58	20,97

### Результаты моделирования

Имитационное моделирование, связанное с формированием на ПЭВМ изображений размером 1024 x 768 пикселей (телевизионная испытательная таблица ТИТ-0249 и таблица ISO 12233) с ПИ, показало, что по сравнению с исходным изображением (рис. 3) при

$$|b_y| < 0,2, |tr_x| < 0,8 \text{ и } |sc_y - 1| < 0,15 \quad (7)$$

разрешающая способность по горизонтали и вертикали ухудшается не более чем на 20 % (рис. 4 и 5).

Сравнение с алгоритмом компенсации дисторсии (2), в котором коэффициенты Русинова вычислялись с использованием метода Левенберга-Маквардта по переопределенной системе из 48 уравнений, проводилось на тестовом изображении размером 1280 x 1024 пикселей типа «сетчатое поле». Результаты эксперимента с диагностической камерой показали, что среднеквадратическое отклонение (СКО) пикселей изображения на АП при формировании ПИ по (4)-(6) и выполнении условий (7) относительно тестового изображения составляет не более 1 пикселя, а при формировании ПИ по (2) – не более 0,4...0,5 пикселя. Несмотря на проигрыш по СКО в 2...2,5 раза ал-

горитм упрощенного формирования ПИ обеспечивает в 2-3 раза меньшие абсолютные искажения на краях проецируемого на комбинер изображения, а также позволяет оперативно (в течение 1-2 минут) изменять параметры компенсации ДИ (например, при изменении базы глаза оператора) без использования диагностической камеры, что важно для практических приложений.

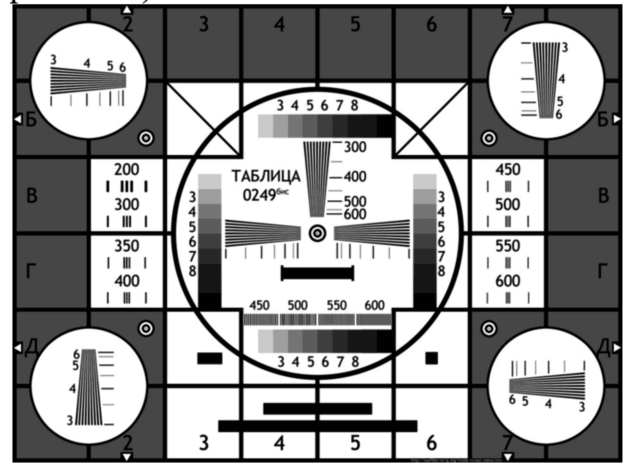


Рис. 3. Тестовая ТВ таблица ТИТ-0249

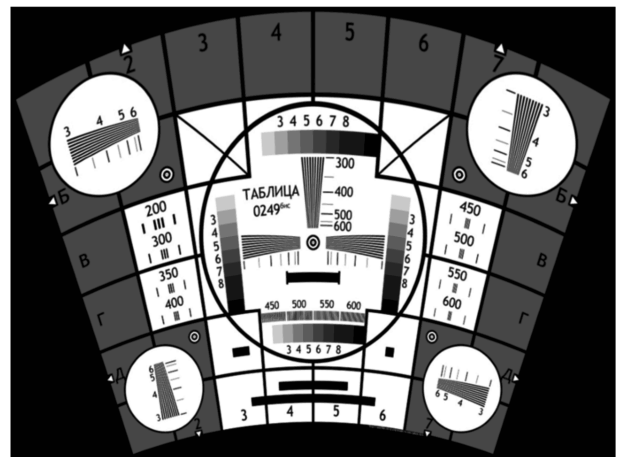


Рис. 4. Кадр видеозображения с ПИ при  $b_y = 0,2, tr_x = -0,8$  и  $sc_y = 1,15$

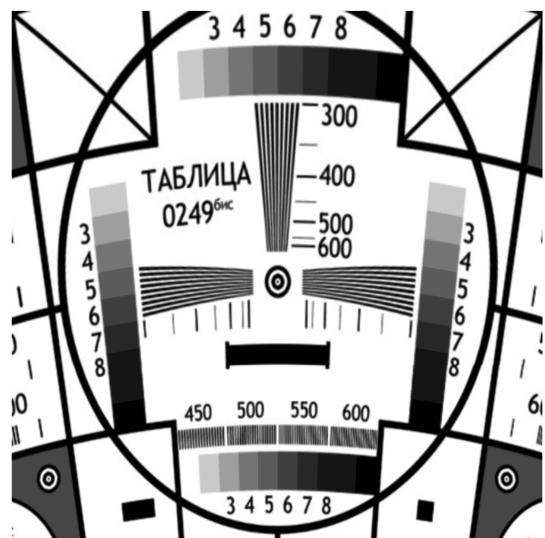


Рис. 5. Центральный фрагмент кадра с рис. 4



## Заключение

Таким образом, рассмотренный алгоритм упрощенного формирования предсказаний позволяет изменением всего трех параметров скорректировать характерные для HUD с асферическими отражающими поверхностями дисторсионные искажения и может быть реализован в реальном масштабе времени.

## Литература

1. Li H. Review and analysis of avionic helmet-mounted displays / H. Li, X. Zhang, G. Shi, H. Qu, Y. Wu, J. Zhang // *Optical Engineering*. – 2013, Vol. 52 (11). – P. 110901-1-110901-14.
2. Козлов А.В. Нашлемная система индикации / А.В. Козлов, И.Г. Денисов, Д.Н. Шарифуллина [Электронный ресурс] // *Будущее машиностроения России: материалы VI Всерос. конф. молодых ученых и специалистов*. – М.: МГТУ им. Н.Э. Баумана, 2013. – 1 эл. опт. диск (CD-ROM).
3. Hua H. Design of an ultra-light head-mounted projective display (HMPD) and its applications in augmented collaborative environments / H. Hua, C. Gao, L. Brown, F. Biocca, J.P. Rolland // *Stereoscopic displays and virtual reality systems. Proceedings of SPIE*. – 2002. – Vol. 4660. – P. 492-497.
4. Бахолдин А.В. Оптические устройства виртуальных дисплеев / А.В. Бахолдин, В.Н. Васильев, В.А. Гримм, Г.Э. Романова, С.А. Смирнов // *Оптический журнал*. – 2013. – № 5. – С. 17-24.
5. Melzer J.E. Head-mounted displays: designing for the user / J.E. Melzer, K.W. Moffitt. – McGraw-Hill, 1997. – 352 p.
6. <http://www.media-phazotron.ru/?p=333>. Проекционные системы отображения вторичной информации авиа- и автобазирования. Дата обращения – 30.06.14.
7. Malacara D. Handbook of optical design, 2nd edition/ D. Malacara, Z Malacara. – New York: Marcel Decker, 2004. – 522 p.
8. Hartley R. Multiple view geometry in computer vision/ R. Hartley, A. Zisserman: 2nd edition. – Cambridge: Cambridge University Press, 2003. – 656 p.
9. Tsai R.Y. A versatile camera calibration technique for high-accuracy 3d machine vision metrology using off-the-shelf TV cameras and lenses / R.Y. Tsai // *IEEE Journal on Robotics and Automation*. – 1987. – RA-3(4). – P. 323-344.
10. Лобанов А.Н. Фотограмметрия / А.Н. Лобанов: 2-е изд., перераб. и доп. – М.: Недра, 1984. – 552 с.
11. Коваленко В.П. Методика оценки дисторсии современных инфракрасных систем / В.П. Коваленко, Ю.Г. Веселов, И.В. Карпиков // *Вестник МГТУ им. Н.Э. Баумана. Сер. «Приборостроение»*. – 2011. – № 1. – С. 98-107.
12. Грузман И.С. Цифровая обработка изображений в информационных системах / И.С. Грузман, В.С. Киричук, В.П. Косых, Г.И. Перетягин, А.А. Спектор. – Новосибирск: Изд-во НГТУ, 2000. – 168 с.
13. Визильтер Ю.В. Обработка и анализ изображений в задачах машинного зрения: курс лекций и практических занятий / Ю.В. Визильтер, С.Ю. Желтов, А.В. Бондаренко, М.В. Ососков, А.В. Моржин. – М.: Физматкнига, 2010. – 672 с.

### SIMPLIFIED DISTORTION COMPENSATION ALGORITHM FOR PROJECTING VIDEO ON PRIORI UNKNOWN FORM ASPHERICAL REFLECTIVE SURFACES

*Kholopov I.S.*

Analytical expressions for simplified forming the predistortion of video frame that projected on priori unknown form aspherical reflecting surfaces is considered.

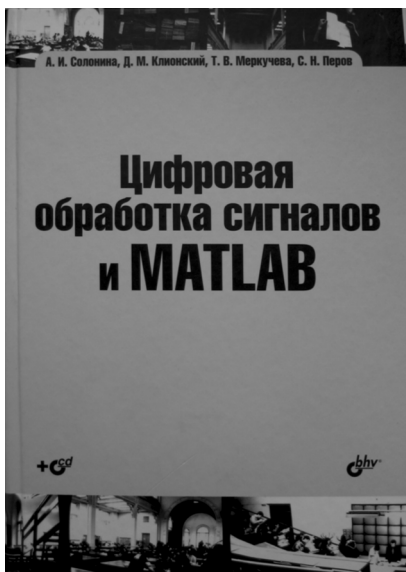
## НОВЫЕ КНИГИ

**Солонина А.И.**

**ЦИФРОВАЯ ОБРАБОТКА СИГНАЛОВ И MATLAB / А.И. Солонина, Д.М. Клинский, Т.В. Меркучева, С.Н. Перов. – СПб.: БХВ-Петербург, 2013. – 512 с. (Учебная литература для вузов)**

**Москва: Техносфера, 2013. – 528 с.**

Описываются базовые методы и алгоритмы цифровой обработки сигналов и средств их компьютерного моделирования в системе MATLAB. Даны основы алгоритмического языка MATLAB. Рассматриваются дискретные сигналы, линейные дискретные системы, дискретное преобразование Фурье с использованием алгоритмов БПФ, синтез и анализ КИХ- и БИХ-фильтров, в том числе с фиксированной точкой, спектральный анализ сигналов, многоскоростная обработка сигналов и адаптивная цифровая фильтрация.



## МЕТОД ВЫДЕЛЕНИЯ КОНТУРОВ ОБЪЕКТОВ С НЕЧЕТКИМИ КРАЯМИ НА ЦВЕТНЫХ ИЗОБРАЖЕНИЯХ

*Медведев Д.Г., начальник компьютерного отдела, ассистент кафедры информатики и прикладной математики Криворожского педагогического института ГВУЗ «Криворожский национальный университет», e-mail: dimakrrog@gmail.com*

**Ключевые слова:** выделение контуров объектов, цветные изображения, размытые границы, сплайн-функции, оптимизация, оператор Соболя.

*Проводится краткий обзор методов выделения контуров объектов. Предлагается метод выделения контуров объектов с размытыми границами с использованием аппроксимирующих сплайн-функций.*

### Введение

Интенсивное развитие средств получения и передачи цифровых изображений создает проблему обработки огромных объемов изображений. Большое количество изображений получают в геологии, минералогии, биологии, металлургии, медицине, метеорологии, экологии и т.п.

Важным источником информации на изображениях является не только наличие или количество определенных объектов, но и их морфология. Например, морфология зерен минералов позволяет определить условия формирования, общности происхождения. Морфология иммунных клеток – эозинофилов – показывает состояние иммунной системы [6]. Однако определение морфологических показателей – трудоемкий процесс, требующий оконтуривания. Эта задача значительно усложняется для микрообъектов, которые по своей природе или из-за особенностей регистрации не имеют четких контуров. Объекты имеют размытые контуры из-за диффузной природы (дым, загрязнения атмосферными выбросами), ввиду «рваного», фрактального характера краев (кроны деревьев, границы больших гранулированных лимфоцитов), из-за разфокусировки изображения.

В оконтуривании наиболее известным и общепринятым является метод активных контуров. Его практический вариант змейка (snake) описан в 1987 году [7]. Функция, моделирующая контур, изменяется, пытаясь достичь минимума собственной энергии (гладкости контура) и несоответствия контура (функция изображения). Метод требует начальной инициализации (как правило 3-5 точек) и в процессе итерационного построения рассчитывает производные и градиенты изображения. Общая идея находит воплощение во многих вариантах метода, отличающихся способами расчета градиентов, производных, методами оптимизации и видом параметрической кривой контура. Разнообразие обусловлено попыткой достичь наибольшей пригодности для конкретных задач, условий и объектов.

Другой подход к оконтуриванию связан с цифровой фильтрацией. Выделение контуров близко к выделению высокочастотной составляющей изображения. Цифровые фильтры выделения контуров реализованы как дискретные дифференциальные операторы 3x3 или 5x5 пикселей. Наиболее известные операторы Роберта, Собеля, Прюитт [4, 5]. Фильтрация изображения этими

операторами позволяет получить контуры изображения во всем поле обработки. Однако по этому изображению трудно определить морфологические показатели отдельных объектов.

Задача оконтуривания значительно усложняется, если контур объекта размыт или недостаточно гладкий. Это приводит к усилению влияния шумов, плохой сходимости итерационного процесса, неточного воспроизведения контуров. Различные авторы, в зависимости от конкретных условий, решают эту задачу по-разному. Например, ведется поиск «поврежденных» фрагментов и настройка на такие фрагменты. Другой метод – метод Hybrid Method использует технологию оконтуривания, при которой выделяются и устраняются по построению контура экстремальные точки. Однако такие алгоритмы требуют дополнительных параметров и их точной настройки, часто при активном участии оператора.

Поэтому задача автоматизации оконтуривания объектов с нечеткими краями за их цифровыми изображениями для определения морфологических параметров и дальнейшей классификации является актуальной.

### Анализ последних исследований и публикаций

Одним из перспективных и современных методов выделения контура объекта на изображении является метод активных контуров. Алгоритм заключается в следующем: производится деформация контура с целью заключения внутри его области, которая содержит пиксели объекта [8].

С помощью метода активных контуров можно найти границу объекта на изображениях, в том случае, если начальная граница изображения задается пользователем или другим способом, возможно автоматизированным. Точки в контуре стремятся к границе объекта при минимизации энергии контура. Для каждой отметки в окрестностях  $v_i$ , значение энергии находится как:

$$E_i = \alpha E_{int}(v_i) + \beta E_{ext}(v_i) \quad (1)$$

где  $E_{int}(v_i)$  является функцией энергии, зависящей от формы контура, и  $E_{ext}(v_i)$  – функцией энергии, зависящей от свойств изображения и типа градиента в окрестности точки  $v_i$ ,  $\alpha$  и  $\beta$  являются константами, обеспечивающими относительную коррекцию величин энергии.

$E_b$ ,  $E_{int}(v_i)$ ,  $E_{ext}(v_i)$  являются матрицами. Значение в центре каждой матрицы соответствует энергии контура в отметке  $v_i$ . Другие значения в матрицах соответствуют энергии в каждой отметке в окрестностях  $v_i$ . Каждая отметка  $v_i$  стремится в  $v_i'$ , соответствующей положению минимального значения в  $E_i$ . Этот процесс иллюстрируется на рис. 1. Если функции энергии выбраны правильно, то контур  $V$  должен приблизиться и остановиться на границе объекта.

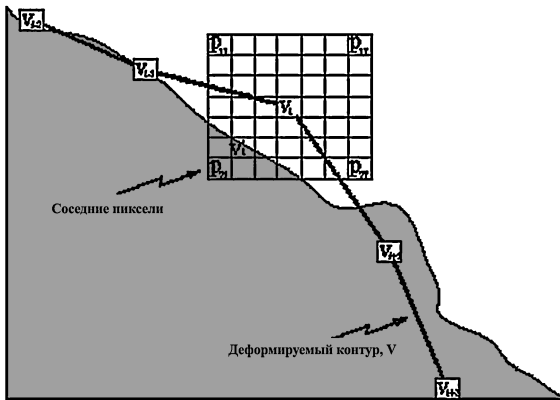


Рис. 1. Пример движения отметки  $v_i$  в активном контуре

Недостаток применения метода в автоматизированных системах состоит в определении коэффициентов  $\alpha$  и  $\beta$ . Они корректируются оператором. Также оператор выбирает и функции энергий внутри контура и за его пределами.

На основе данного метода разработаны алгоритмы для выделения контуров объектов различной природы. Быстрый алгоритм активных контуров, предложенный Donna J. Williams и Mubarak Shah [4], выделяет контуры различных объектов, однако, как и в самом методе, необходимо прямое участие оператора при выборе коэффициентов.

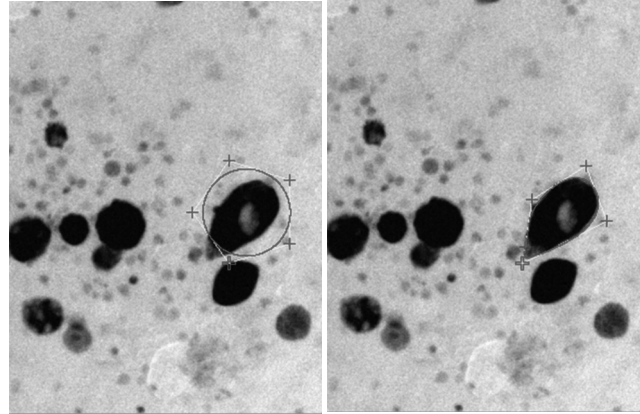
Компанией Biomedical Image Group (BIG) [2] предложен метод моделирования контуров биологических объектов методом змейки (E-Snake) с применением экспоненциальных сплайнов (рис. 2) [3].

Такой алгоритм является наиболее подходящим для морфологического анализа клеток.

Среди различных возможных представлений кривой используется B-сплайн, который характеризуется  $M$  точками управления. Точки (крестики) задают контрольный полигон (прямые линии между точками), который управляет сплайн-змейкой (кривая линия).

Эволюция кривой сформулирована как задача минимизации. Связанная с ней функция называется энергией змейки. Используются комбинации контура и области энергий. Энергия контура соответствует интегралу от проекции изображения градиента на кривой. Эта энергия использует сглаженную версию исходного изображения с гауссовским ядром  $\alpha$  стандартного отклонения. Области энергий максимизируют разницу между интенсивностью области, ограниченной кривой, и интенсивностью вне ее.

Как отмечают авторы, экспоненциальный сплайн в B-форме лучше всего подходит для моделирования круговых, эллиптических и каплеобразных форм. При аппроксимации минимизируют разность энергий области, ограниченной кривой, и выделенной области. Для минимизации этого показателя используется итерационная процедура оптимизации контрольных точек, определяющих полигон, описанный вокруг объекта моделирования.



а) до оптимизации б) после оптимизации

Рис. 2 Выделенный контур клетки экспоненциальным сплайном

Однако процедура оптимизации несколько осложнена тем, что контрольные точки движутся на плоскости. Точки не имеют прямой связи со строящимся контуром.

При использовании частотных фильтров для выделения контуров чаще всего пользуются операторами Робертса, Собеля, Прюитт. Операторы Собеля и Прюитт используют градиент, вычисленный по формуле  $G = \sqrt{G_x^2 + G_y^2}$ . Отличие состоит лишь в способе вычисления производных по составляющим. Оператор Робертса использует окно  $2 \times 2$ , и величина градиента вычисляется по формуле:

$$G_R(j, k) = \sqrt{[F(j, k) - F(j+1, k+1)]^2 + [F(j, k+1) - F(j+1, k)]^2} \quad (2)$$

Вместе с тем, при использовании данных методов выделяются контуры всех объектов на изображении. Количество найденных контуров зависит от выбранного порогового значения. Если его выбрать высоким, то получается мало контуров, и искомый объект может быть невыделен. При низком значении – большое количество контуров. То есть, в общем случае необходимо участие оператора для выбора порогового значения, что усложняет процесс автоматизации выделения контура.

Автором были выполнены попытки применить данные методы к оконтуриванию объектов с размытыми контурами, но результаты оказались неудовлетворительными (рис. 3).

**Цель исследований** – разработать метод выделения контура объектов с нечеткими, размытыми краями на цветных изображениях, лишенный в той или иной степени отмеченных выше недостатков.

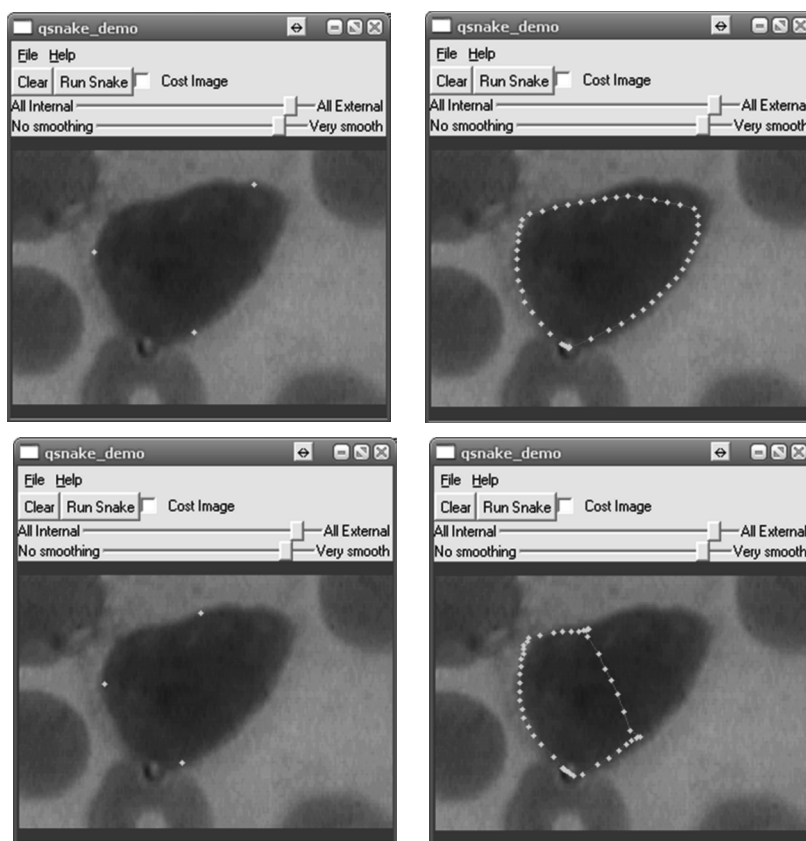


Рис. 3. Применение Qsnake к оконтуриванию эозинофила

### Результаты исследований и разработки

В работе предлагается информационная технология оконтуривания объектов с использованием метода, состоящего из последовательных этапов.

Целью предварительной обработки является улучшение качества получаемого изображения. Снимки, как правило, недостаточно четкие и контрастные, а также несут в себе много лишней информации, прежде всего шумов (пыль, осевшую на стекла мазков, другие факторы). Улучшить четкость и контрастность можно аппаратными методами, например, заменой видеокамеры на ту, которая специализируется на съемке изображений под микроскопом. Но это дорогостоящий способ, который не решает проблему исключения шумов. Поэтому целесообразнее воспользоваться алгоритмическими методами улучшения качества изображения.

Уменьшение изолированных препятствий с минимальным размыванием контура обеспечивается за счет одномерной медианной фильтрации по пяти точкам. Перепады яркости на изображении содержат в себе информацию о контурах объектов. Именно с их помощью выполняется разделение объектов друг от друга. Перепады яркости обычно разделяют друг от друга две области изображения с различной яркостью. Идеальный детектор перепада должен указывать наличие перепадов в единой точке, расположенной строго в центре градиента яркости. Для грубого выделения контура применен оператор Собеля с окном 3\*3, элементы которого показаны на рис 4.

Расчеты выполняются по формуле:

$$G(j,k) = \sqrt{X^2 + Y^2}, \quad (3)$$

где  $X = (A_2 + 2A_3 + A_4) - (A_0 + 2A_7 + A_6)$ ,

$Y = (A_0 + 2A_1 + A_2) - (A_6 + 2A_5 + A_4)$

$A_0$	$A_1$	$A_2$
$A_7$	$F(j,k)$	$A_3$
$A_6$	$A_5$	$A_4$

Рис. 4. Окно расчетов оператором Собеля

Однако, в общем случае, это не позволяет достичь оконтуривания именно необходимого объекта. Кроме того, оператор Собеля требует выбора порогового значения. Для работы оператора изображение переводится в формат YUV. Оператор применяется к сигналу яркости Y.

После использования этих методов получаем границы объектов – бинарное изображение, но границы находятся неоднозначно, как множество отдельных точек. Последнее объясняется нечеткими и размытыми границами тела.

Учитывая только яркость, невозможно выделить нужный контур и определить его важные характеристики. Поэтому необходимо принимать во внимание и компоненты цветности U (сине-желтая) и V (красно-голубая). Для этого к бинарному изображению яркости (компонента Y) добавим компоненты цветности в неизменном виде. Получаем цветное изображение с другим распределением цветов за счет бинарного преобразования яркости. На нем четко выделены те же контуры, но объект стал белого цвета, а фон имеет желтоватый цвет.

Благодаря разнице в цвете выделяем объект. Для этого разработана стратегия последовательного перебора. Находятся приблизительные границы «белого тела». Находятся белые точки, вокруг которых есть доста-

точное количество (не менее 5) таких же белых точек. Это исключает отдельные белые точки среди цветного фона. Далее двигаемся в четырех направлениях до черных точек. В результате получаем декартовы координаты точек контура искомого объекта и прямоугольник, в который объект вписан. Это очень грубое выделение контура, по которому сложно определить морфологические параметры.

Для точного оконтуривания объекта, представленного множеством отдельных точек, предложен новый метод итерационного приближения контура сплайном. Особенность искомого контура в том, что он является замкнутой функцией. Построить циклический, замкнутый сплайн можно двумя способами: переходом к параметрической функции или полярной системе координат.

При использовании параметрической функции появляется сложность выбора параметра  $t$ , и имеем две отдельные функции  $x(t)$  и  $y(t)$ . Поэтому был избран перевод в полярную систему координат.

Исследуемые объекты (микроорганизмы, минералы) имеют форму, близкую к выпуклой, что упрощает переход, как будет показано далее на рис. 8-10.

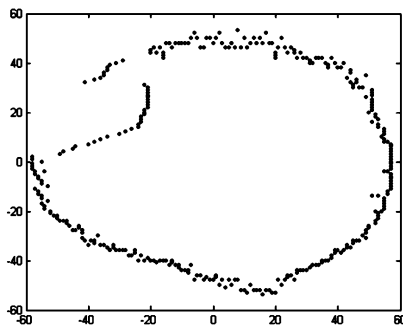
Началом отсчета полярной системы координат служит центр прямоугольника, в который вписан найденный контур (рис. 5).

Для уточнения грубо выделенного контура необходимо построить сплайн-модель. Особенность данного сплайна в том, что он является замкнутым. В таком сплайне количество узлов равно количеству фрагментов. Особенностью периодического сплайна является то, что перед первым фрагментом находится последний, а за последним фрагментом следует первый. Для построения сплайн-модели был выбран сплайн Эрмита. Для вычислений базисных функций сплайна Эрмита используются следующие формулы [2]:

$$\begin{aligned} X_0(x) &= -p_1 p_4 b, & X_1(x) &= p_1 c - p_2 d + b, \\ X_2(x) &= -p_4 c - p_3 d + a, & X_3(x) &= -p_3 p_2 d, \\ a &= (x - x_j) / h_j, & b &= 1 - a, & c &= ab^2, & d &= a^2 b, \\ p_1 &= \frac{h_j}{h_{j-1}}, & p_2 &= \frac{h_j}{h_{j+1} + h_j}, & p_3 &= \frac{h_j}{h_{j+1}}, & p_4 &= \frac{h_j}{h_{j-1} + h_j}, \\ h_{j-1} &= x_j - x_{j-1}, & h_j &= x_{j+1} - x_j, & h_{j+1} &= x_{j+2} - x_{j+1}. \end{aligned} \quad (4)$$

Замкнутость контура сплайна учитывается следующим образом:

- для первого:  $h_{j-1} = x_R - x_{R-1}$ ;
- для последнего:  $h_j = x_j - x_0, h_{j+1} = x_2 - x_1$ ;
- для предпоследнего:  $h_{j+1} = x_l - x_0$ ,



где  $R$  – количество фрагментов сплайна,  $j$  – номер текущего фрагмента.

В общем виде, для сплайна, который состоит из  $R$  фрагментов, имеем выражение:

$$S(x) = \sum_{j=0}^R f(x_j) X_j(x) \quad (5)$$

где  $X_j(x)$  локальная базисная сплайн-функция, состоящая из четырех фрагментов:

$$X_j(x) = \begin{cases} X_{0,j-1}(x), & x \in [x_{j-1}, x_j); \\ X_{1,j}(x), & x \in [x_j, x_{j+1}); \\ X_{2,j+1}(x), & x \in [x_{j+1}, x_{j+2}); \\ X_{3,j+2}(x), & x \in [x_{j+2}, x_{j+3}); \\ 0, & x \notin [x_{j-1}, x_{j+3}). \end{cases} \quad (6)$$

Значение сплайна в произвольной точке, принадлежащей  $j$ -му фрагменту, равно:

$$S(x) = f(x_{j-1})X_{0,j-1}(x) + f(x_j)X_{1,j}(x) + f(x_{j+1})X_{2,j+1}(x) + f(x_{j+2})X_{3,j+2}(x) \quad (7)$$

Для замкнутого сплайна это выражение истинно для всех фрагментов.

Аппроксимация выполняется по методу наименьших квадратов (МНК). Необходимо на множестве точек  $y_i, a_i, i = \overline{1, N}$ , оценить истинное значение контура  $f(a_i)$  при условии аддитивной погрешности с нулевым математическим ожиданием  $y_i = f_i + \varepsilon_i, M[\varepsilon_i] = 0$ . Погрешность считаем некоррелированной и распределенной по нормальному закону с дисперсией  $D[\varepsilon] = \sigma^2$ . Необходимо найти такие значения вектора  $A = \{a_{ij}\}$ , при которых достигается минимум  $d = \sum_{i=1}^N (y_i - S(a_i, A))^2 \rightarrow \min$ .

В общем виде МНК с моделью сплайном не отличается от аналогичных уравнений для алгебраических полиномов.

Решением является:

$$A = (P^T P)^{-1} (P^T Y) \quad (8)$$

где  $P$  – матрица планирования, столбцы – базисы сплайна (6).

Особенностью сплайн-модели контура является его замкнутость. Начало и конец сплайна являются условными, так как конечный узел одновременно является и первым. Поэтому матрица планирования для замкнутого сплайна имеет блочно-диагональную форму с дополнительными блоками справа-вверху и слева-внизу (рис. 6).

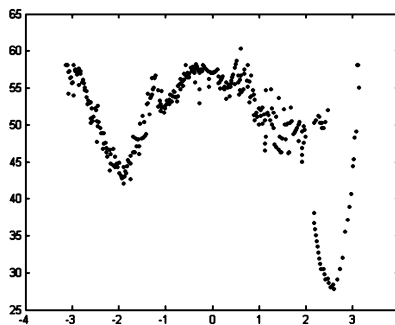


Рис. 5. Функция контура в декартовой и полярной системе координат

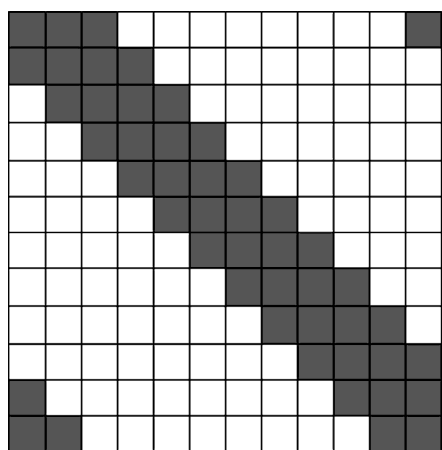


Рис. 6 Схема расположения блоков матрицы планирования

Для построения матрицы планирования  $P$  используются формулы (4). Для каждой  $i$ -ой точки вычисляются значения функций  $X_0(x_i)$ ,  $X_1(x_i)$ ,  $X_2(x_i)$ ,  $X_3(x_i)$ . В матрице планирования они находятся в  $i$ -ой строчке. Для каждого  $j$ -го фрагмента с общего количества фрагментов  $R$ , кроме  $j = 1, R - 1, R$ , значения функций  $X_0(x_i)$ ,  $X_1(x_i)$ ,  $X_2(x_i)$ ,  $X_3(x_i)$  находятся в столбцах с номерами  $j - 1, j, j + 1, j + 2$  соответственно. Для первого фрагмента ( $j = 1$ ) значения функций  $X_0(x_i)$  записываются в столбец  $R$ , так как в качестве нулевого фрагмента используется последний. Значения  $X_1(x_i)$ ,  $X_2(x_i)$ ,  $X_3(x_i)$  записываются в столбцы  $1, 2, 3$  соответственно. Для предпоследнего фрагмента ( $j = R - 1$ ) значения  $X_3(x_i)$  заносятся в  $1$ -ый столбец, так как за последним фрагментом идет первый. Значения  $X_0(x_i)$ ,  $X_1(x_i)$ ,  $X_2(x_i)$  записываются в столбцы  $R - 2, R - 1, R$  соответственно. По той же причине для последнего фрагмента ( $j = R$ ) значения  $X_3(x_i)$  заносятся во  $2$ -й столбец,  $X_2(x_i)$  заносятся в  $1$ -ый столбец. Значения  $X_0(x_i)$ ,  $X_1(x_i)$ , записываются в столбцы  $R - 1, R$  соответственно.

Гладкость и точность приближения сплайн-контура неявно зависит от количества узлов сплайна. Для достижения приближения сплайном в  $\sigma < 3$  разработан метод оптимизации количества узлов и их размещения. На начальной стадии сплайн строится на четырех фрагментах с четырьмя узлами. Первый узел находится в первой точке сплайна. Остальные три – на равных расстояниях друг от друга. Уточнение положения узлов сплайна выполняется следующим образом. Расстояние между соседними узлами делится на пять частей. Каждый узел движется в каждую из них. Если среднеквадратическое отклонение сплайн контура после временного движения узла уменьшилась, то отныне узел находится в этой новой точке. Если после оптимизации размещения узлов не удается достичь необходимой точности приближения, то увеличиваем число фрагментов сплайна (добавляем еще один узел): в середину фрагмента с наибольшим отклонением добавляется новый узел. Далее вновь повторяется оптимизация размещения узлов. В среднем сплайн строится на 10-12 узлах. Результат оптимизации показан на рис. 7.

Разработанный метод можно применять при оконтуривании объектов различной природы. С его помощью были выделены контуры на цифровых изображениях минераграфии (рис. 8), металлографии (рис. 9), иммунологии (рис. 10).

Предложенный метод реализован в информационной технологии определения морфологических параметров эозинофилов. Исследуя морфологические признаки больших гранулированных лимфоцитов (эозинофилов), можно определить иммунный статус человека, то есть стадию борьбы организма с вирусами. Морфологические признаки эозинофилов здорового и больного человека существенно отличаются. Для определения иммунного статуса пациента применен классический дискриминантный анализ [8].

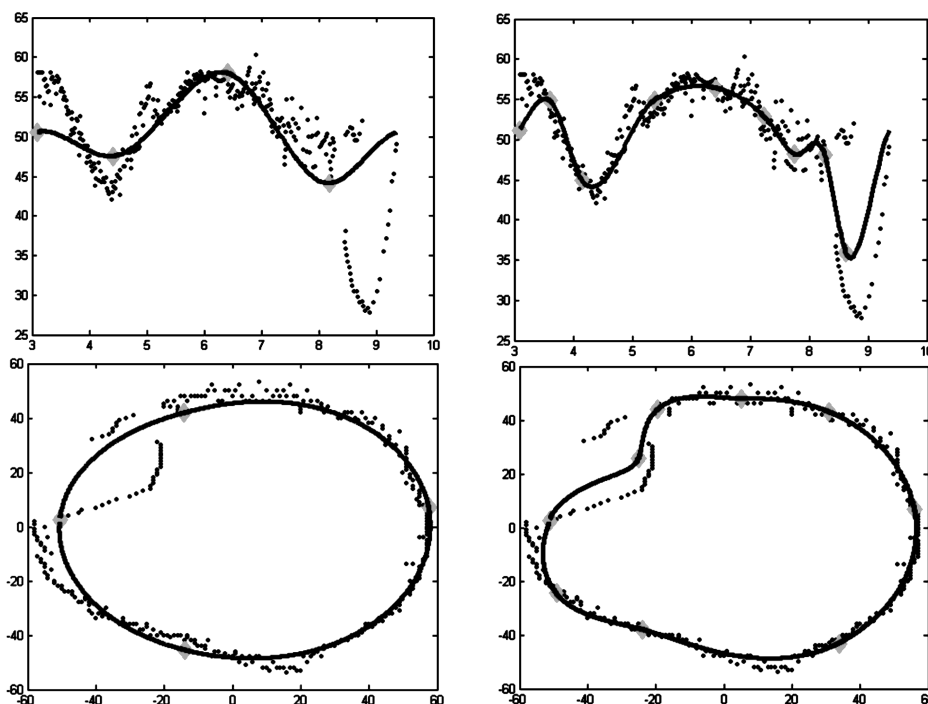


Рис. 7. Контур объекта до и после оптимизации

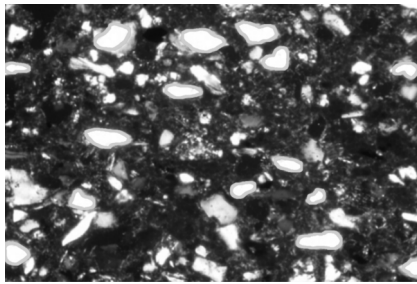


Рис. 8. Контуры срезов минералов

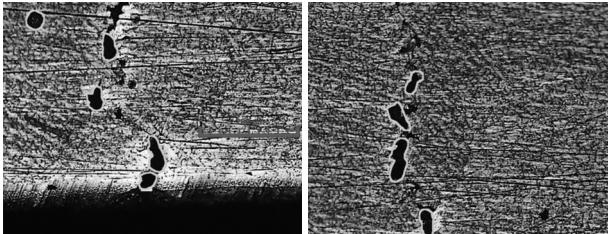


Рис. 9. Контуры поврежденной металлической поверхности

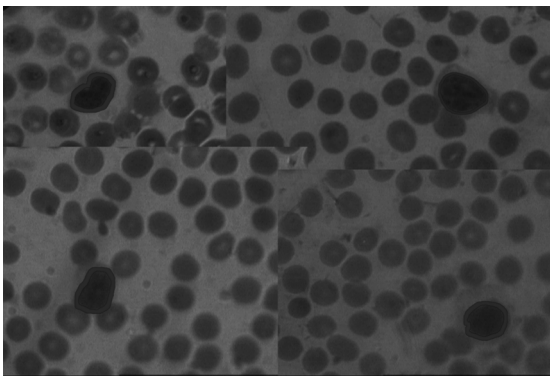


Рис. 10. Контуры объектов микроскопии

Для определения дискриминантной функции использованы относительные показатели:  $x_1 = \frac{d_v}{d_m}$ ,

$$x_2 = \frac{S}{l^2 / 4\pi}, \quad x_3 = \frac{l}{2\sqrt{\pi S}},$$

где  $d_v$  – большой диаметр,  $d_m$  – малый диаметр,  $S$  – площадь,  $l$  – длина контура. Для расчета использована обучающая выборка из  $n = 110$  образцов, из них  $n_1 = 72$ ,  $n_2 = 38$ .

Определены коэффициенты  $\beta_i$  дискриминантной функции и получена дискриминантная функция  $Y(X) = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$ .

Вероятность ошибочной классификации определяется следующим образом. Если  $m_1$  – число наблюдений  $W_1$ , отнесенных к  $W_2$ , и  $m_2$  – число наблюдений  $W_2$ , классифицированных в  $W_1$ , то  $P(2|1) = m_1/n_1$  и  $P(1|2) = m_2/n_2$ .

По данным ретроспективного анализа было получено  $P(2|1) = 0,073$ , а  $P(1|2) = 0,126$ . Это вероятность отношения здорового пациента к больным, а больного – к здоровым.

### Заключение

Разработанный метод, работа которого проверялась на реальных изображениях, продемонстрировал устойчивость и достижение заданной точности.

Метод включает в себя последовательность следующих операций.

1. Преобразование изображения в YUV.
2. Медианная фильтрация.
3. Определение грубого контура оператором Собеля.
4. Определение координат контура.
5. Аппроксимация сплайном.
6. Оптимизация контура.

Небольшое нарушение выпуклости объектов не является препятствием для предложенного метода, хотя в отдельных случаях может несколько снижать точность за счет сглаживания.

### Литература

1. Шелевицкий І.В., Шутко М.О., Шутко В.М., Калганова О.О., Сплайни в цифровій обробці даних і сигналів. Кривий Ріг: Видавничий дім. 2008. – 232с., іл.
2. Biomedical Imaging Group: Gradient (index) <http://bigwww.epfl.ch/algorithms.html>
3. DONNA J. WILLIAMS AND MUBARAK SHAH A Fast Algorithm for Active Contours and Curvature Estimation. CVGIP: IMAGE UNDERSTANDING Vol. 55, No. 1, January, pp. 14-26, 1992
4. Roberts L.G. Machine Perception Of Three-Dimensional Solids / L.G. Roberts – Massachusetts Institute of Technology Lincoln Laboratory. Режим доступа: <http://www.packet.cc/files/mach-per-3D-solids.html>
5. Samuel J. Dwyer III. A personalized view of the history of PACS in the USA. / Samuel J. Dwyer III. // Proceedings of the SPIE, «Medical Imaging 2000: PACS Design and Evaluation: Engineering and Clinical Issues». Edited by G. James Blaine and Eliot L. Siegel, 2000. – P. 2-9.
6. Литвинова Т.В. Профілактика респіраторних захворювань у дітей, хворих на бронхіальну астму, що одержують базисну терапію: дис. канд. мед. наук: 14.01.10 / Т.В. Литвинова. – Д.:ДДМА, – 2006. – 193с.
7. Kaas M., Witkin A., Terzopoulos D. Snakes: Active Contour Models. // Int. Journal of Computer Vision. – 1987, N1, p.312-331.
8. Факторный, дискриминантный и кластерный анализ / Дж.-О. Ким, Ч.У. Мьюллер, У.Р. Келка и др.; под ред. И.С. Енюкова. Пер. с англ.– М.: Финансы и статистика, 1989. – С. 90-93.

### THE METHOD OF EDGE DETECTION OF OBJECTS WITH FUZZY EDGES IN COLOR IMAGES

*Medvedev D. G.*

A brief review of methods of edge detection of objects is considered. Basic methods do not allow to allocate the contours of objects with blurred edges or require the active participation of the user. The advantage of the proposed method of edge detection using the approximating spline functions is its automatization (the user only need to specify one point belonging to the object). The result is a contour in the form of the analytic function that simplifies parameterization. The developed method can be used in edge detection of objects of different nature: mineragraphy, metallo-graphy, immunology. The method is implemented in the information technology for determination of morphological parameters of eosinophils.

## НЕЙРОСЕТЕВОЙ АЛГОРИТМ УТОЧНЕНИЯ ВЕКТОРНЫХ ТОПОГРАФИЧЕСКИХ КАРТ ПО ДАННЫМ ДИСТАНЦИОННОГО ЗОНДИРОВАНИЯ ЗЕМЛИ

*Акинин М.В., аспирант кафедры «Космические технологии» Рязанского государственного радиотехнического университета, e-mail: akinin.m.v@gmail.com*

**Ключевые слова:** искусственная нейронная сеть, машина опорных векторов, нейронная карта Кохонена, многослойный перцептрон, нейронная сеть прямого распространения грамма графовой генерации Китано, генетический алгоритм.

### Введение. Постановка задачи

Задача оперативного составления и обновления топографических карт является одной из важнейших решаемых задач в различных сферах экономики, природопользования и обеспечении безопасности жизнедеятельности человека.

Данная задача может быть решена посредством анализа данных дистанционного зондирования Земли (ДЗЗ) полу- и полностью автоматическими интеллектуальными системами, что обеспечивает высокую точность результата и низкие временные затраты на его получение.

Топографическая карта подготавливается в геоинформационной системе на основе набора векторных слоев, каждый из которых описывает положение различных объектов местности. Задача уточнения топографической карты сводится, таким образом, к задаче уточнения набора векторных слоев, которая, в свою очередь, сводится к получению векторного слоя  $V$  для каждого исходного слоя:

$$V_{old} \xrightarrow{f} V,$$

где  $f$  – функционал уточнения, имеющий вид

$$V = f(V_{old}, I).$$

Здесь  $I$  – спутниковый снимок или данные аэрофотосъемки.

$I$  можно представить в виде

$$I = \{i(m, n, k)\}; i(m, n, k) \in \{-1, 1\};$$

$$m = \overline{1, M}, n = \overline{1, N}, k = \overline{1, K}; M, N, K \in \mathbb{N},$$

где  $i(m, n, k)$  – значение в спектральном канале  $k$  пикселя с координатами  $(m, n)$ ;  $M, N, K$  – количество строк, столбцов и спектральных каналов в изображении  $I$ .

Каждому пикселю  $i(m, n)$  можно поставить в соответствие набор координат  $(x, y, z)$  в системе координат карты:  $(x, y, z) = \text{map}(m, n)$ . Таким образом достигается наложение данных ДЗЗ на векторные слои карты.

Предметом рассмотрения в настоящей статье явля-

*Рассмотрен алгоритм уточнения векторных топографических карт по данным дистанционного зондирования Земли. Предложено использование интеллектуальной системы, основанной на машинах опорных векторов, многослойных перцептронах, нейронных сетях прямого распространения, нейронной карте Кохонена. Рассмотрен подход к обучению нейронной сети прямого распространения с помощью генетического алгоритма и грамматик графовой генерации Китано.*

ется функционал  $f$  уточнения векторных слоев карты. Он может быть построен разными методами, однако в настоящее время все чаще стали использоваться методы искусственного интеллекта, основанные на теории машинного обучения. В настоящей статье рассматривается только простейший – двухмерный случай. Описываемый алгоритм может применяться также и для уточнения трехмерных моделей местности.

### Алгоритм уточнения

Предлагаемый алгоритм уточнения топографических карт состоит из следующих этапов.

**Этап 1.** Предварительная обработка данных ДЗЗ  $I_{raw}$  – коррекция шумов, конвертирование в систему координат карты, проецирование в двухмерную проекцию карты (координата  $z$  здесь и далее принимается равной нулю).

**Этап 2.** Преобразование  $I_{raw}$  из исходного формата представления спектральных каналов к формату представления, используемому алгоритмом:

$$i(m, n, k) = \text{spectral\_convert}(i_{raw}(m, n, k)) =$$

$$= a * i_{raw}(m, n, k) + b;$$

$$m = \overline{1, M}, n = \overline{1, N}, k = \overline{1, K}; a, b \in \mathbb{R},$$

где  $a, b$  – коэффициенты линейного преобразования, выбираемые в зависимости от исходного формата представления спектральных каналов.

Так как чаще всего исходным форматом является беззнаковое целое 8-ми битное число, то коэффициенты  $a, b$  принимают значения:

$$a = \frac{2}{255}, b = -1.$$

**Этап 3.** Классификация пикселей изображения  $I$  с получением на выходе маски  $C$ :

$$c(m, n) = \text{classification}(i(m, n), V_{old}, I_{old});$$

$$c(m, n) \in P;$$

$$m = \overline{1, M}; n = \overline{1, N}; P \subset \mathbb{N},$$



где  $P$  – множество индексов классов. Каждый индекс  $p$  маркирует один из возможных классов объектов, которые могут присутствовать в векторном слое,  $I_{old}$  – данные ДЗЗ, по которым был построен векторный слой  $V_{old}$ .

**Этап 4.** Векторизация маски  $C$  с получением на выходе векторного слоя  $V$ . Элементы  $c(m, n)$  маски  $C$  составляют элемент векторного слоя  $V$  (точку, линию или полигон – в зависимости от типа слоя) в случае, если они имеют одинаковое значение (принадлежат к одному и тому же классу) и образуют связное множество точек в декартовой системе координат относительно центра маски.

Векторизация может быть выполнена с помощью алгоритмов, описанных в [1] и [2].

### Классификатор

Ключевым этапом представленного алгоритма является классификация пикселей изображения  $I$  с получением на выходе маски  $C$ . Классификация выполняется с помощью классификатора *classification*. Существуют различные способы построения классификатора: ручные, полуавтоматические (с ручным подбором параметров алгоритма, с ручным этапом верификации результатов классификации и прочие) и автоматические. Наибольший интерес представляют автоматические способы построения классификатора из-за возможности существенно снизить временные затраты на классификацию.

В качестве автоматически настраиваемого классификатора *classification* могут быть использованы решения, основанные на концепциях теории искусственного интеллекта и теории машинного обучения.

В ходе подготовки статьи был разработан классификатор, работающий по алгоритму, состоящему из следующих этапов.

**Этап 1.** Предварительная классификация пикселей  $i(m, n)$  с формированием на выходе предварительной маски  $C_{pre}$ . Для предварительной классификации пикселей используется нейронная карта Кохонена.

**Этап 2.** Сопоставление маски  $C_{pre}$  и векторного слоя  $V_{old}$  с формированием на выходе списка классифицированных и неклассифицированных пикселей  $i(m, n)$ .

Каждый пиксель  $i(m, n)$  маркируется как классифицированный в случае, если мера расстояния  $d(I(m, n), I_{old}(m, n))$  не превосходит некоторого порога  $T$ ;  $T \in \mathbb{R}^+$ , выбираемого адаптивно в зависимости от гистограмм изображений  $I, I_{old}$  – чем меньше контраст изображений (чем ярче выражены моды гистограммы и чем ближе моды гистограммы друг к другу), тем меньше значение порога  $T$ , что необходимо для повышения чувствительности алгоритма на однородных изображениях. Метрика  $d$  выбирается в зависимости от решаемой задачи. Как показали эксперименты, для уточнения топографических карт разумным выбором является

Евклидова мера расстояния.

**Этап 3.** Обучение ядра классификатора *kernel*. Ядро классификатора обучается на классифицированных пикселях  $i(m, n)$ .

**Этап 4.** Классификация пикселей  $i(m, n)$  с формированием на выходе маски  $C$ . Если пиксель  $i(m, n)$  включен в список классифицированных, то  $c(m, n) = c_{pre}(m, n)$ , иначе  $c(m, n) = kernel(i(m, n))$ .

### Ядро классификатора

Ядро классификатора представляет собой набор простых классификаторов *classifier\_simple*, организованных в древовидную структуру.

Каждый простой классификатор выполняет разбиение исходного векторного подпространства  $Q_{sub} \subset Q$  векторного пространства  $Q$ , составленного из векторов  $i(m, n)$ , на два класса. Будучи отнесенным к тому или иному классу, вектор  $i(m, n)$  спускается на следующий уровень дерева до тех пор, пока не достигнет одного из стоков. Каждый из стоков дерева соответствует одному из результирующих индексов  $p$  возможных классов.

Обучение ядра классификатора сводится к обучению простых классификаторов.

Обучение простого классификатора является итеративным процессом, состоящим из следующих этапов:

- подбор параметров обучения и прогона классификатора;
- обучение классификатора по тому или иному алгоритму.

В качестве простых классификаторов могут быть использованы следующие концепции теории машинного обучения:

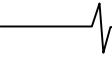
- машина опорных векторов [4];
- многослойный перцептрон [5];
- нейронная сеть прямого распространения.

### Нейронная сеть прямого распространения

Интеллектуальная нейронная сеть (ИНС) прямого распространения может быть представлена как функция  $y = ann(v)$  – функция, производящая для каждого входного вектора  $v$  соответствующий отклик  $y$ .

ИНС прямого распространения состоит из нескольких нейронов, связанных друг с другом синапсами, по которым передаются сигналы, циркулирующие в сети. Каждый нейрон обладает некоторым тормозящим пороговым значением, определяющим уровень реакции нейрона на соответствующее возбуждение. Термин «прямое распространение» означает, что в ИНС отсутствуют циклы. Каждый нейрон сети подвергает свой входной вектор некоторому преобразованию в соответствии со своим пороговым значением и своей функцией активации, передавая свой выход по соответствующим синапсам другим нейронам.

Большинство ИНС прямого распространения относятся к классу нейронных сетей, обучаемых с учителем. Для обучения таковой ИНС используется обучающее



множество  $V_{teach}$  векторов, для которого известно множество идеальных откликов  $d_{teach}$ . Для тестирования качества обучения ИНС используется тестовое множество  $V_{test}$  векторов, для которых также известны идеальные отклики  $d_{test}$ .

В общем случае, подбор структуры ИНС прямого распространения представляется сложной задачей. Существует ряд подходов к ее решению, из которых наиболее развитым является подход с использованием генетического алгоритма [6] для поиска оптимальной структуры нейронной сети с точки зрения среднеквадратичного функционала ошибки

$$E(ann, V_{test}) = \frac{\sum_i (ann(v_{test\ i}) - d_{test\ i})^2}{V_{test}}; \quad i = \overline{1, V_{test}}$$

Важной проблемой, решаемой в процессе поиска оптимальной структуры ИНС с помощью генетического алгоритма, является проблема кодирования структуры ИНС в виде непрерывной последовательности символов – последовательности хромосом, составляющих единственный ген генома каждой особи.

Н. Kitano в [7] предложил способ кодирования структуры ИНС, называемый грамматиками графовой генерации Китано. В настоящей статье рассматривается одно из развитий грамматик графовой генерации Китано, которое заключается в введении дополнительного параметра  $W$ , определяющего глубину разбора начального правила грамматики и позволяющего задавать с помощью грамматик, содержащих одно и то же количество правил, сети произвольного размера.

Структура ИНС прямого распространения может быть представлена в виде ориентированного графа без циклов. Каждая дуга в данном графе соответствует синапсу и имеет собственный вес, устанавливаемый в процессе обучения сети. Каждая вершина графа соответствует одному нейрону; с каждой вершиной связано пороговое значение соответствующего нейрона, устанавливаемое в процессе обучения сети.

Имеющиеся в графе дуги могут быть заданы с помощью матрицы смежности  $G$  (1) ИНС, в которой единица кодирует наличие ребра, направленного от вершины с номером, равным номеру строки, к вершине с номером, равным номеру столбца, а ноль – соответственно, отсутствие такого ребра.

$$G = \begin{pmatrix} g_{11} & \dots & g_{1H_G} \\ \dots & \dots & \dots \\ g_{H_G 1} & \dots & g_{H_G H_G} \end{pmatrix};$$

$$g_{ij} \in \{0, 1\}; \quad i = \overline{1, H_G}; \quad j = \overline{1, H_G}.$$

Матрицу  $G$  можно закодировать с помощью набора правил порождающей грамматики  $A$ :

$$A = \{T, T_p, T_N, F_T, F_p, F_N, S\};$$

$$T = \{0, 1, \omega\}; \quad T_p = \{T_p\ i\}; \quad i = \overline{1, 16};$$

$$T_N = \{T_N\ j\}; \quad j = \overline{1, H_N}; \quad S \in T_N;$$

$$F_T = \{1 \rightarrow \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}; \quad 0 \rightarrow \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}; \quad \omega \rightarrow 0\};$$

$$F_p = \{F_{p1} \rightarrow \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad F_{p2} \rightarrow \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \dots, F_{p16} \rightarrow \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\};$$

$$F_N = \{F_{N\ i}\}; \quad F_{N\ i}: T_{N\ i} \rightarrow \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix};$$

$$i = \overline{1, H_N}; \quad \alpha, \beta, \gamma, \delta \in (F_N \cup F_p),$$

где:  $T$ ,  $T_p$ ,  $T_N$ ,  $F_T$ ,  $F_p$ ,  $F_N$ ,  $S$  – наборы терминальных символов, предтерминальных символов, нетерминальных символов, правил вывода для терминальных символов, правил вывода для предтерминальных символов, правил вывода для нетерминальных символов и стартовый нетерминальный символ соответственно.

Грамматика  $A$  называется грамматикой графовой генерации Китано. Одна грамматика  $A$  соответствует одной особи популяции генетического алгоритма. Элементы  $T$ ,  $T_p$ ,  $T_N$ ,  $F_T$ ,  $F_p$  фиксированы для любой  $A$ , и, следовательно, ген особи кодирует только набор правил  $F_N$ . Набор правил  $F_N$  может быть закодирован по схеме, приведенной на рис. 1. Каждое правило  $F_{Ni}$  отображается в четыре хромосомы  $\alpha_i$ ,  $\beta_i$ ,  $\gamma_i$ ,  $\delta_i$ , следующие друг за другом и за хромосомой  $\delta_{i-1}\sigma_X^2$  в единственном гене  $b$  особи.

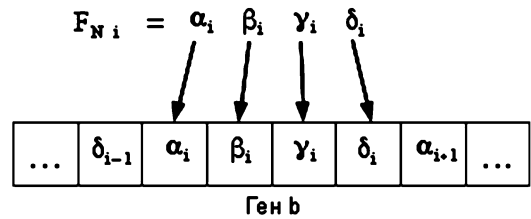


Рис. 1. Схема кодировки набора правил  $F_N$

Алгоритм восстановления матрицы  $G$  из грамматики  $A$  состоит из следующих этапов.

**Этап 1.** Создание начальной матрицы  $G_1 = S$ .

**Этап 2.** Выполнение  $w=1$  шаг применения правил  $F_T$ ,  $F_p$  и  $F_N$ :  $G_1 = S \rightarrow G_2 = \begin{pmatrix} \alpha_S & \beta_S \\ \gamma_S & \delta_S \end{pmatrix}$ .

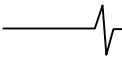
**Этап 3.** Повторение этапа 2 для каждого из правил  $\alpha_S$ ,  $\beta_S$ ,  $\gamma_S$ ,  $\delta_S$ :  $G_2 = \begin{pmatrix} \alpha_S & \beta_S \\ \gamma_S & \delta_S \end{pmatrix} \rightarrow G_3$ ; размерность  $G_3$

оказывается, таким образом, равной  $2^{(3-1)}$  на  $2^{(3-1)} = 4$  на 4 правила.

**Этап 4.** Повторение этапа 3 ( $W-2$ ) раз - до получения матрицы  $G_W$ .

**Этап 5.** В матрице  $G_W$  любой символ  $g \in (T_p \cup T_N)$  рассматривается как символ  $\omega$  и преобразуется по соответствующему правилу в 0.

**Этап 6.** Преобразование матрицы  $G_W$  в матрицу  $G$  по формуле (2) – таким образом, в матрице  $G$  гарантированно будут отсутствовать циклы.



$$G = (g_{ij}); \quad G_W = (g_{W \ ij});$$

$$g_{ij} = \begin{cases} 0, & i \geq j \\ g_{W \ ij}, & i < j \end{cases}; \quad i = \overline{1, H_G}; \quad j = \overline{1, H_G} \quad (2)$$

Параметр  $W$  алгоритма восстановления матрицы смежности  $G$  определяет максимально возможное количество нейронов, которое может быть в результирующей ИНС –  $2^{W-1}$  нейронов. Учитывая тот факт, что число входов и выходов ИНС фиксировано, то к ИНС, описываемой соответствующим графом, присоединяются несколько нейронов:

– входные нейроны в достаточном количестве – связываются синапсами с истоками графа; данные синапсы направлены от входных нейронов к истокам графа;

– выходные нейроны в достаточном количестве – связываются синапсами со стоками графа; данные синапсы направлены от стоков графа к выходным нейронам.

Мощность  $H_N$  множества  $T_N$  оказывает, равно как и параметр  $W$ , существенное влияние на качество работы генетического алгоритма, поскольку именно оно определяет длину гена (как  $4H_N$  хромосом) и, соответственно, степень «разнообразия» правил, кодируемых геном. Выбор  $H_N$  осуществляется эвристическим способом – небольшое значение  $H_N$  приведет к вырождению популяции и ее сходимости к локальному оптимуму, тогда как больше значение  $H_N$ , наоборот, увеличит разнообразие особей, но также увеличит время схождения генетического алгоритма к, возможно, глобальному оптимуму.

По сравнению с традиционной версией грамматики графовой генерации Китано, предложенная версия позволяет одинаково компактно описывать нейронные сети с любым максимальным количеством нейронов, что достигается за счет параметра  $W$ .

### Генетический алгоритм

Генетический алгоритм, используемый для поиска оптимальной структуры ИНС и использующий для кодирования структуры ИНС грамматики графовой генерации Китано, состоит из следующих этапов.

**Этап 1.** Создание множества  $V = \{v_i\}$ ;  $i = \overline{1, H_V}$  векторов из исследуемой предметной области и множество  $d = \{d_i\}$ ;  $i = \overline{1, H_V}$  идеальных откликов ИНС на соответствующие вектора из  $V$ ; на основе множеств  $V$  и  $d$  на каждой итерации генетического алгоритма будут составляться обучающие  $V_{teach}$ ,  $d_{teach}$  и тестовые  $V_{test}$ ,  $d_{test}$  множества векторов и откликов.

**Этап 2.** Генерация случайной начальной популяции особей  $B = \{b_i\}$ ;  $i = \overline{1, H_B}$ .

**Этап 3.** Выделение из множества  $V$  случайным образом подмножеств  $V_{teach}$ ,  $V_{test}$ ; для множеств  $V_{teach}$  и  $V_{test}$  из множества  $d$  выделяются соответствующие им подмножества  $d_{teach}$  и  $d_{test}$ .

**Этап 4.** Восстановление грамматики  $A_i$  для каждой

особи  $b_i \in B$ .

**Этап 5.** Расчет для каждой особи  $b_i \in B$  функции приспособленности

$$Q(b_i) = \frac{1}{E(ann_i, V_{test}) + \varepsilon}; \quad \varepsilon \in \mathbb{R}; \quad \varepsilon < 10^{-8}.$$

**Этап 6.** Сортировка порядка следования особей  $b_i \in B$  по убыванию функции приспособленности  $Q(b_i)$ .

**Этап 7.** Особи с индексами  $i \in [1, H_{B \ best}]$  образуют множество особей  $B_{best} \subset B$ , особи с индексами  $i \in (H_{B \ best}, H_{B \ best} + H_{B \ cross}]$  – множество особей  $B_{cross} \subset B$ ; особи с индексами  $i \in (H_{B \ best} + H_{B \ cross}, H_{B \ best} + H_{B \ cross} + H_{B \ mut}]$  – множество особей  $B_{mut} \subset B$ , при этом должны выполняться условия:

$$H_{B \ best} \geq 2;$$

$$H_{B \ cross} = \sigma H_{B \ mut}; \quad \sigma \in \mathbb{R}; \quad \sigma > 1;$$

$$H_{B \ best} + H_{B \ cross} + H_{B \ mut} = H_B;$$

$$B_{best} \cap B_{cross} = \emptyset; \quad B_{best} \cap B_{mut} = \emptyset; \quad B_{cross} \cap B_{mut} = \emptyset.$$

**Этап 8.** Замена особей  $b_i \in B_{cross}$  на особи  $b_{cross \ kl} = cross(b_k, b_l)$ ;  $b_k, b_l \in B_{best}$  – то есть на особи – результаты скрещиваний случайных особей из множества  $B_{best}$  лучших особей популяции.

**Этап 9.** Замена особей  $b_i \in B_{mut}$  на особи  $b_{mut \ k} = mutation(b_k)$ ;  $b_k \in B_{best}$  – то есть на особи – результаты мутаций случайных особей из множества  $B_{best}$  лучших особей популяции.

**Этап 10.** Если лучшая особь популяции  $b_i$  была лучшей и соответствовала условию останова  $Q(b_i) > P$  на менее чем  $I$  предыдущих итераций алгоритма подряд, то выполняется переход на шаг 3, иначе выполняется останов алгоритма. Параметр  $P$  задает уровень, определяющий достаточное качество обучения ИНС.

По останову выполнения генетического алгоритма ИНС, генерируемая грамматикой  $A_i$ , соответствующей особи  $b_i$  окончательной популяции, принимается как лучшая.

Генетический алгоритм обладает следующими параметрами:

–  $H_V$ ;  $H_{V \ teach} = \{V_{teach}\}$ ;  $H_{V \ test} = \{V_{test}\}$  – определяют размеры обучающей и тестовой выборки векторов; значения данных параметров зависят от предметной области и сложности задачи классификации, решаемой искомым ИНС;

–  $H_{B \ best}$ ;  $H_{B \ cross}$ ;  $H_{B \ mut}$ ;  $\sigma$  – определяют количество особей, сохраняемых при переходе к следующей итерации генетического алгоритма, заменяемых на результаты скрещивания и мутации соответственно;

–  $cross$ ,  $mutation$  – функционалы скрещивания и мутации соответственно;

–  $P$  – определяет верхний уровень качества обуче-

ния лучшей ИНС, достаточный для останова алгоритма;  
–  $I$  – определяет достаточное для останова алгоритма количество итераций, на которых подряд должны повторяться условия останова для лучшей особи; очевидно, что чем больше  $I$ , тем больше вероятность схождения алгоритма к устойчивому оптимуму – к ИНС, дающей одинаковый по качеству обучения результат на произвольных  $V_{teach}$  и  $V_{test}$ .

### Эксперимент

Для проверки состоятельности предлагаемого способа уточнения топографических карт было поставлено несколько экспериментов, один из которых будет рассмотрен далее.

Эксперимент заключается в уточнении и дополнении топографической карты полигона «Малинищи», расположенного на границе Рязанского и Пронского районов Рязанской области и названного так по названию села Малинищи, расположенного на данном полигоне.

В качестве уточняемых данных ДЗЗ  $I_{old}$  был взят спутниковый снимок части полигона «Малинищи», сделанный сенсором «ETM +» КА «Landsat – 7» 22-го мая 2000-го года. На основе  $I_{old}$  была создана уточняемая топографическая карта  $\{V_{old}\}$ . Задача разработанного алгоритма состоит в том, чтобы уточнить исходную топографическую карту части полигона «Малинищи» и дополнить результирующую карту информацией обо всем полигоне. Уточнение и дополнение топографической карты было выполнено по данным  $I$  от 23-го мая 2006-го года, полученным тем же сенсором.

По сравнению с топографической картой  $\{V_{old}\}$ , составленной для  $I$  вручную, алгоритм продемонстрировал высокую точность результата, неправильно классифицировав менее 5% пикселей  $i(m, n)$ .

В среднем, точность алгоритма достигает 93% правильно классифицированных пикселей  $i(m, n)$  по сравнению с ручной классификацией.

Разработанный алгоритм показал высокую временную эффективность – спутниковый снимок  $I$  с размерами  $M = 4000$ ,  $N = 4000$ ,  $K = 6$  может быть обработан в среднем за 30 – 40 мс, что достаточно для обработки 25 кадров в секунду (25 FPS (англ. frame per second – количество кадров в секунду) – частота, достаточная для формирования видеопоследовательности, плавной и непрерывной для зрительной системы среднестатистического человека).

### Заключение

В ходе проведения экспериментов разработанный алгоритм продемонстрировал следующие достоинства:

– высокая точность результата – не менее 93 % пикселей исходных данных ДЗЗ классифицируются правильно;

– низкие временные затраты – 30–40 мс. на 6-ти канальное изображение размером 4000 на 4000 пикселей (типичный спутниковый снимок), что позволяет достичь 25 FPS;

– высокий уровень автоматизации процесса уточнения.

Недостатками разработанного алгоритма являются высокая сложность его реализации и необходимость предварительной автоматической настройки классификатора для улучшения его временных характеристик.

### Литература

1. Шапиро Л., Стокман Д. Компьютерное зрение – М.: БИНОМ. Лаборатория знаний. - 2006.

2. Гонсалес Р., Вудс Р. Цифровая обработка изображений // Пер. с английского под редакцией П.А. Чочиа. – М.: Техносфера. – 2005.

3. Хайкин С. Нейронные сети: Полный курс, 2-е изд., испр.: Пер. с англ. – М.: ООО «И.Д. Вильямс» – 2006.

4. Акинин М.В., Конкин Ю.В. Технология тематического дешифрирования спутниковых изображений на основе машины опорных векторов. // Информатика и прикладная математика: межвузовский сборник научных трудов. – Рязань: Рязанский государственный университет. – 2010.

5. Акинин М.В., Конкин Ю.В. Исследование подходов к обучению многослойного перцептрона. // Методы и средства обработки и хранения информации: межвузовский сборник научных трудов. – Рязань: Рязанский государственный радиотехнический университет. – 2012.

6. Аксенов С.В., Новосельцев В.Б. Организация и использование нейронных сетей (методы и технологии). – Томск: НТЛ. – 2006. – 128 с.

7. Kitano H. Designing neural network using genetic algorithm with graph generation system // Complex Systems. Vol. 4. – 1990.

### USING ARTIFICIAL NEURAL NETWORK TO CREATE TOPOGRAPHIC MAP BASED ON REMOTE SENSING DATA

*Akinin M.V.*

Paper describes algorithm to create topographic vector map based on remote sensing data. Paper describes intellectual system based on support vector machine, multilayer perceptron, feed forward neural network and Kohonen's self-organized map. Paper describes algorithm to teach feed forward neural network using genetic algorithm and Kitano's graph generation grammatic.