

ЦИФРОВАЯ ОБРАБОТКА СИГНАЛОВ

3/2009

НАУЧНО - ТЕХНИЧЕСКИЙ ЖУРНАЛ

ЦИФРОВАЯ ОБРАБОТКА ИЗОБРАЖЕНИЙ

тематический выпуск:

Нам 10 лет!



- КОДИРОВАНИЕ И СЖАТИЕ
- СКРЫТИЕ ИНФОРМАЦИИ
- НОРМАЛИЗАЦИЯ ИЗОБРАЖЕНИЯ
- ФОРМИРОВАНИЕ СНИМКОВ
- СЛЕЖЕНИЕ ЗА ОБЪЕКТАМИ
- ВИДЕОПРОЦЕССОР ГЛАЗА

ISSN 1684-2634



9 771684 263005 >

ЦИФРОВАЯ ОБРАБОТКА СИГНАЛОВ

Научно-технический журнал
№ 3/2009

Издается с 1999 года
Выходит четыре раза в год

ГЛАВНЫЙ РЕДАКТОР
Ю.Б. ЗУБАРЕВ

ЗАМЕСТИТЕЛИ ГЛАВНОГО РЕДАКТОРА:
В.В. ВИТЯЗЕВ, В.П. ДВОРКОВИЧ

РЕДАКЦИОННАЯ КОЛЛЕГИЯ:

Б.А. Бабаян, Ю.И. Борисов,
С.А. Грибачев, Г.В. Зайцев, Р.В. Зубарев,
А.П. Кирпичников, М.И. Кривошеев,
Н.А. Кузнецов, М.С. Куприянов, А.А. Ланнэ,
В.Г. Мистюков, С.Л. Мищенко,
А.А.Петровский, Ю.Н. Прохоров,
А.Н. Соловьев, Ю.Г. Сосулин,
В.В. Шахильдян, Ю.С. Шинаков

Адрес редакции:

107031 Москва, Рождественка, 6/9/20, стр. 1
Тел.: (495) 621-71-08, (495) 621-16-39
Факс: (495) 621-16-39
E-mail: nto.popov@mtu-net.ru
tor@rgta.ryazan.ru
http://www.dspsa.ru

Для писем:

129090 Москва, а/я 48.

Издатель:

Российское научно-техническое общество радиотехники,
электроники и связи им. А.С. Попова
Ответственный редактор: П.И. Зудков
Компьютерная верстка: В.В. Корягина
Дизайн: Н.С. Осипова

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

Подписано в печать 12.10.09
Формат 60x90/8.

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

Заказ № 5385. 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

УЧРЕДИТЕЛИ:

ЗАО «Инструментальные системы»
ФГУП «НИИ радио»

В НОМЕРЕ:

- Умняшкин С.В., Курина В.В.**
Алгоритм сжатия изображений на основе дискретного псевдосинусного преобразования 2
- Богуш Р.П.**
Стеганографическое скрывание информации в статических изображениях 8
- Медведева Е.В.**
Метод выделения контуров в изображении на основе вычисления количества информации 12
- Давыденко Е.В., Приоров А.Л.**
Калибровка положения видеокамеры в задаче оптической лазерной триангуляции 16
- Егошкин Н.А., Еремеев В.В., Козлов Е.П.**
Нормализация космических изображений Земли на основе их сопоставления с электронными картами 21
- Еремеев В.В., Князьков П.А., Москвитин А.Э.**
Оценка разрешающей способности аэрокосмических изображений на основе их статистического анализа ... 27
- Еремеев В.В., Зенин В.А.**
Радиометрическая нормализация изображений от многоматричных сканирующих устройств 31
- Кузнецов А.Е., Светелкин П.Н.**
Формирование цветных снимков по спутниковым изображениям среднего и высокого пространственного разрешения 36
- Кузнецов А.Е., Пошехонов В.И.**
Оценка высоты облачности по данным комплекса мезозональной сканерной съемки космического аппарата "Метеор - М" 41
- Гулевич А.Е., Дворкович А.В., Кочарян А.Э., Мингазов И.Д.**
Реализация многофункционального кодера AVC 47
- Алпатов Б.А., Бабаян П.В., Смирнов С.А.**
Автоматическое слежение за объектами при отсутствии априорных сведений о фоноцелевой обстановке 52
- Кирпичников А.П.**
Основной видеопроцессор нашей цивилизации – уникальный пример эволюции в новейшей истории? 57
- Витязев С.В.**
Учебная лаборатория ЦОС Рязанского государственного радиотехнического университета 63

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

АЛГОРИТМ СЖАТИЯ ИЗОБРАЖЕНИЙ НА ОСНОВЕ ДИСКРЕТНОГО ПСЕВДОКОСИНУСНОГО ПРЕОБРАЗОВАНИЯ

Умняшкин С.В., д.ф.-м.н., профессор Московского государственного института электронной техники (технический университет)

Курина В.В., Московский государственный институт электронной техники (технический университет), valya_kurina@rambler.ru

Ключевые слова: сжатие изображений, алгоритм, дискретное преобразование, матрица, квантование, кодирование.

Введение

В настоящее время известно множество быстрых алгоритмов вычисления дискретного косинусного преобразования (ДКП), разработанных для различных задач сжатия изображений и видео. Они позволяют сократить количество операций умножения, необходимых для выполнения преобразований, но, тем не менее, эта операция не может быть полностью исключена [1]. Если требуется выполнять вычисления без операций умножения, можно воспользоваться преобразованиями, приближенными к ДКП, например [2, 3].

В данной работе мы предлагаем использовать для сжатия изображений дискретное псевдокосинусное преобразование [4] – более простую в вычислении альтернативу ДКП, не требующую операций умножения. Алгоритм сжатия изображений на основе такого преобразования может быть востребован в устройствах с ограниченными аппаратными ресурсами, например, карманных компьютерах или мобильных телефонах.

Дискретное псевдокосинусное преобразование

Матрица дискретного псевдокосинусного преобразования (ДПКП) \mathbf{W} может быть представлена в виде произведения двух матриц специального вида: $\mathbf{W} = \mathbf{D} \cdot \mathbf{C}$, где \mathbf{D} – диагональная нормировочная матрица, а \mathbf{C} – матрица, допускающая факторизованное представление в виде слабозаполненных матриц.

Так, матрица восьмиточечного ДПКП имеет вид [4]: $\mathbf{W}_8 = \mathbf{D}_8 \cdot \mathbf{C}_8$, где

$$\mathbf{D}_8 = \text{diag} \left(\frac{1}{2\sqrt{2}}, \frac{1}{2\sqrt{5}}, \frac{1}{2\sqrt{2}}, \frac{1}{2\sqrt{5}}, \frac{1}{2\sqrt{2}}, \frac{1}{2\sqrt{5}}, \frac{1}{2\sqrt{2}}, \frac{1}{2\sqrt{5}} \right),$$

$$\mathbf{C}_8 = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 2 & 1 & -1 & -2 & 2 & 1 & -1 & -2 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & -2 & 2 & -1 & 1 & -2 & 2 & -1 \\ \hline 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 2 & 1 & -1 & -2 & -2 & -1 & 1 & 2 \\ 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 \\ 1 & -2 & 2 & -1 & -1 & 2 & -2 & 1 \end{bmatrix}.$$

При этом, за счет использования специального факторизованного представления, умножение на матрицу \mathbf{C}_8 сводится к выполнению 28 операций сложения чисел [4].

В общем случае, при других размерностях ДПКП, реализация умножения на матрицу \mathbf{C} также сводится только к

Предлагается алгоритм сжатия статических полутоновых изображений, который основан на быстром целочисленном преобразовании, полученном в результате модификации дискретного косинусного преобразования.

выполнению операций сложения. ДПКП, так же, как и ДКП, является ортогональным преобразованием.

Для оценки эффективности применения ортогональных преобразований с целью сжатия сигнала можно воспользоваться величиной *coding gain* (коэффициент эффективности кодирования) [5]:

$$C_g = 10 \lg \frac{\frac{1}{N} \sum_{k=0}^{N-1} \sigma_k^2}{\left(\prod_{k=0}^{N-1} \sigma_k^2 \right)^{\frac{1}{N}}}, \quad (1)$$

где σ_k^2 – дисперсия k -го коэффициента преобразования.

Значения коэффициента C_g для ДПУА – дискретного преобразования Уолша-Адамара, ДПКП, ДКП и ДПКЛ – дискретного преобразования Карунена-Лозва приведены в Таблице 1. Расчеты выполнены для входного сигнала ($N = 8$), имеющего статистику марковского процесса первого порядка с коэффициентом корреляции соседних отсчетов $\rho = 0,95$.

Таблица 1.

Преобразование	ДПУА	ДПКП	ДКП	ДПКЛ
C_g	7,9461	8,3251	8,8259	8,8462

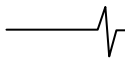
Таким образом, с одной стороны ДПКП не намного уступает ДКП по эффективности применения к сжатию данных, с другой – требует меньших вычислительных затрат, так как не использует операции умножения.

Базовый алгоритм сжатия

Предлагаемая нами базовая «JPEG-подобная» схема компрессии статического полутонного изображения приведена на рис 1. В процессе декомпрессии последовательно выполняются операции арифметического декодирования, скалярного деквантования и двумерного обратного ДПКП. Исходное изображение разбивается на блоки размера 8×8 пикселей. На этапе вычисления двумерного ДПКП сначала строки каждого блока изображения, затем столбцы полученной матрицы умножаются на матрицу \mathbf{C}_8 . Равномерное скалярное квантование спектрального коэффициента ДПКП w объединяется с умножением на матрицу $\mathbf{D}_8 = \{d_{i,j}\}$ благодаря введению специального шага квантования

$$q_{i,j} = \frac{q}{d_{i,i} d_{j,j}},$$

где индексы i и j определяются положением квантуемого коэффициента в блоке. Квантование выполняется в со-



ответствии с правилом:

$$\hat{w} = \text{Round}(w/q_{i,j}), \quad (2)$$

где функция Round производит округление до ближайшего целого числа. Деквантование значения \hat{w} выполняется по правилу:

$$\tilde{w} = q_{i,j} \hat{w}. \quad (3)$$

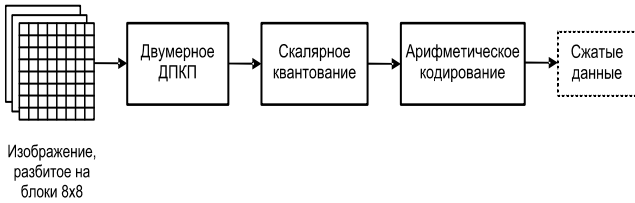


Рис. 1. Схема компрессии

Пусть $w_{i,j}^{(l)}$ – коэффициент с номером $l = 0, \dots, 63$ в блоке с координатами (i, j) (начало координат – блок в левом верхнем углу). Нумерация внутри блока – от левого верхнего угла по строкам. Поток данных из постоянных составляющих каждого блока (коэффициентов $w_{i,j}^{(0)}$) после квантования подвергается двумерной дифференциальной импульсно-кодовой модуляции (ДИКМ). Вычисление разности двумерной ДИКМ можно осуществить с помощью операций сложения и побитового сдвига в соответствии со следующим соотношением:

$$d_{i,j} = \hat{w}_{i,j}^{(0)} - \frac{1}{2} (2\hat{w}_{i-1,j}^{(0)} + 2\hat{w}_{i,j-1}^{(0)} - \hat{w}_{i-1,j-1}^{(0)} - \hat{w}_{i-1,j+1}^{(0)}). \quad (4)$$

Для реализации арифметического кодирования был выбран адаптивный кодер [6, 7]. Характерной особенностью этого кодера является использование целочисленной арифметики, однако в процессе вычислений требуется выполнение операций умножения и деления. В случае необходимости дополнительного сокращения (устранения) операций умножения, можно воспользоваться реализациями арифметических кодеров из работ [8, 9], в которых вычисления выполняются с помощью операций сложения и побитового сдвига, но при этом снижается эффективность кодирования [8].

Спектральные коэффициенты ДПКП в каждом блоке распределяются на восемь зон, кодируемых с помощью разных статистических моделей распределения вероятностей (рис. 2а). Зоны выбирались таким образом, чтобы среднеквадратичные отклонения (СКО) коэффициентов были близки, и наиболее коррелированные коэффициенты входили в одну зону. Данные о СКО спектральных коэффициентов приведены на рис. 2б. С целью сокращения объема алфавитов кодирования моделей предлагается кодировать модули коэффициентов, а знаки ненулевых коэффициентов записывать в отдельный поток с затратами 1 бит на символ.

Зону, в которой все коэффициенты равны нулю, будем называть нулевой. Для описания нулевых зон каждому блоку поставим в соответствие восьмибитовое число – ключ блока. Обозначим через $k_{i,j}$ значение ключа блока с координатами (i, j) , $k_{i,j}(r)$ – значение r -го бита числа $k_{i,j}$, $Z_{i,j}^r$ – множество проквантованных спектральных коэффициентов, образующих зону с номером r в блоке с координатами (i, j) , тогда

$$k_{i,j}(r) = \begin{cases} 0, & \text{если } \forall \hat{w}_{i,j}^{(l)} \in Z_{i,j}^r : \hat{w}_{i,j}^{(l)} = 0, \\ 1, & \text{если } \exists \hat{w}_{i,j}^{(l)} \in Z_{i,j}^r : \hat{w}_{i,j}^{(l)} \neq 0. \end{cases} \quad (5)$$

Таким образом, на этапе адаптивного арифметического кодирования формируются и кодируются независимо следующие потоки данных: поток ключей блоков (кодируемый с помощью 1 статистической модели), поток модулей разностей от двумерной ДИКМ постоянных составляющих блоков (1 модель), поток модулей спектральных коэффициентов из ненулевых зон (8 моделей).

Алгоритм обработки блока спектральных коэффициентов с координатами (i, j) на этапе арифметического кодирования состоит из следующих шагов:

1. Арифметическое кодирование постоянной составляющей. Вычисление разности $d_{i,j}$ по формуле (4). Абсолютное значение разности $|d_{i,j}|$ кодируется в отдельном потоке. Если $d_{i,j} \neq 0$, то знак $s_{i,j} = \text{sign}(d_{i,j})$ (1 бит) записывается в поток знаков.

	0	3	3	0	0	3	3
1	5	5	7	2	6	6	7
4	5	7	7	6	6	7	7
4	7	7	7	7	7	7	7
1	2	6	7	2	2	5	7
1	6	6	7	2	2	5	7
4	6	7	7	5	5	7	7
4	7	7	7	7	7	7	7

а)

	23,3	8,41	4,55	44,9	23,0	8,86	4,25
14,2	5,43	3,93	2,34	9,01	6,87	3,79	2,45
3,97	3,47	3,10	1,96	3,93	3,87	2,81	1,98
2,73	1,91	1,76	1,45	2,20	1,97	1,71	1,49
27,8	9,92	6,05	3,08	19,0	13,6	6,52	3,37
12,2	7,23	5,30	2,74	11,9	10,8	5,37	2,81
4,61	3,45	2,92	1,89	4,71	4,46	2,60	1,96
2,30	1,96	1,86	1,54	2,17	2,06	1,70	1,56

б)

Рис. 2. Расположение зон в блоке (а) и СКО коэффициентов блока двумерного ДПКП (б)



2. Формирование ключа блока $k_{i,j}$ по правилу (5) и его арифметическое кодирование.

3. Арифметическое кодирование спектральных коэффициентов из ненулевых зон. Просматриваем последовательно все коэффициенты блока (за исключением постоянной составляющей), $l=1, \dots, 63$. Пусть $\widehat{w}_{i,j}^{(l)} \in Z_{i,j}^r$, тогда, если $k_{i,j}(r)=1$, закодировать $|\widehat{w}_{i,j}^{(l)}|$, используя модель с номером r , и, если $\widehat{w}_{i,j}^{(l)} \neq 0$, то записать $s_{i,j}^{(l)} = \text{sign}(\widehat{w}_{i,j}^{(l)})$ в файл знаков. Если $k_{i,j}(r)=0$, то коэффициент $\widehat{w}_{i,j}^{(l)}$ не обрабатывается.

Модификация базового алгоритма

Квантование с нулевой зоной

В базовом варианте алгоритма сжатия выполнялось равномерное скалярное квантование с порогами $\{\dots, -2,5q, -1,5q, -0,5q, 0,5q, 1,5q, 2,5q, \dots\}$ и уровнями – в середине интервалов. То есть спектральные коэффициенты $w_{i,j}^{(l)}$, такие, что $-0,5q < w_{i,j}^{(l)} \leq 0,5q$, после квантования принимали значение, равное 0.

При квантовании с нулевой зоной (dead zone quantization) [10] изменяется интервал, соответствующий нулевому уровню. Пороги квантования принимают вид: $\{\dots -3kq, -2kq, -kq, kq, 2kq, 3kq, \dots\}$, а уровни квантования вычисляются как среднее арифметическое соседних порогов. Значение параметра k выбирается так, чтобы большая часть высокочастотных компонент квантовалась в 0. Такое квантование используется, например, в стандартах JPEG2000, MPEG-2, H.263.

Результаты обработки нескольких тестовых изображений для нашей схемы компрессии показали, что оптимальным является значение $k=1$. Таким образом, в окончательном варианте квантование выполнялось с порогами $\{\dots, -3q, -2q, -q, q, 2q, 3q, \dots\}$.

Многомодельное арифметическое кодирование

Как видно из рисунка 2б, в зонах коэффициентов, кодируемых при помощи одной и той же статистической модели, все же присутствуют коэффициенты, имеющие заметно различающиеся СКО. Для более точного моделирования распределения вероятностей коэффициентов каждой зоны можно использовать не одну, а несколько моделей, выбирая необходимую по «контексту», то есть на основании некоторого правила, использующего результаты предшествующей работы арифметического кодера.

В нашей схеме компрессии мы использовали по две модели для кодирования проквантованных коэффициентов каждой зоны рис. 2а. Правило выбора модели кодирования для текущего обрабатываемого коэффициента $\widehat{w}_{i,j}^{(l)}$ строилось следующим образом:

$$\text{Model}(\widehat{w}_{i,j}^{(l)}) = \begin{cases} 2r, & \text{Prog}(\widehat{w}_{i,j}^{(l)}) \leq t_r \\ 2r+1, & \text{Prog}(\widehat{w}_{i,j}^{(l)}) > t_r \end{cases} \quad (6)$$

Здесь r – номер зоны, в которой содержится коэффициент $\widehat{w}_{i,j}^{(l)}$, t_r – порог для выбора модели кодирования, соответствующий зоне r , $\text{Prog}(\widehat{w}_{i,j}^{(l)})$ – прогноз, представляющий собой некоторую функцию от уже обработанных «контекстных» коэффициентов ДПКП.

В алгоритмах сжатия изображений на базе ДКП возможны различные способы выбора контекста: на осно-

вании внутривыделочной [11, 12] и межвыделочной [13, 14] взаимной корреляции коэффициентов спектра. В алгоритме на основе ДПКП мы ограничились построением прогноза по коэффициентам из соседних с кодируемым блоком, которые соответствуют одинаковым базисным функциям:

$$\text{Prog}(\widehat{w}_{i,j}^{(l)}) = \frac{1}{6} \left(2|\widehat{w}_{i-1,j}^{(l)}| + 2|\widehat{w}_{i,j-1}^{(l)}| + |\widehat{w}_{i-1,j-1}^{(l)}| + |\widehat{w}_{i-1,j+1}^{(l)}| \right). \quad (7)$$

Весовые коэффициенты в данной формуле учитывают тот факт, что корреляция кодируемого коэффициента с соответствующими коэффициентами из блоков сверху и слева от текущего блока (i, j) больше, чем с коэффициентами из «диагональных» блоков $(i-1, j-1)$ и $(i-1, j+1)$. Заметим, что при реализации кодера деление на 6 можно опустить, увеличив в 6 раз значения порогов $\{t_r\}$ в решающем правиле (6). Таким образом, вычисление прогноза (7) вновь сводится только к выполнению операций сложения.

По результатам обработки широко распространенных тестовых изображений «Lena», «Barbara» и «Goldhill» для различных значений шага квантования q были эмпирически найдены оптимальные величины порогов $\{t_r\}$ и установлено, что их значения достаточно хорошо описываются формулой:

$$t_r = c_r / q, \quad (8)$$

где c_r – постоянные масштабные коэффициенты, определенные на основании экспериментальных данных по методу наименьших квадратов. Таким образом, пороги для правила (6) являются варьируемыми и выбираются по значению используемой величины шага равномерно квантования q при помощи формулы (8), а набор значений $\{c_r\}$ является фиксированным.

RD-оптимизация

Дополнительно повысить характеристики алгоритма сжатия можно, воспользовавшись идеями теории «битовые затраты - искажение» (rate-distortion theory, RD theory) [15].

Поставим следующую задачу RD-оптимизации: необходимо найти такой набор ключей блоков $\mathbf{k}^* = \{k_{11}^*, k_{12}^*, \dots, k_{nm}^*\}$ и такую величину шага квантования q^* , которые бы обеспечивали минимум функции Лагранжа:

$$L(\mathbf{k}^*, q^*) = \min_{\mathbf{k}, q} (L(\mathbf{k}, q) = D(\mathbf{k}, q) + \lambda R(\mathbf{k}, q)). \quad (9)$$

Здесь \mathbf{k} – всевозможные наборы ключей блоков; λ – задаваемый внешне неотрицательный параметр, который определяет соотношение качества изображения и битовых затрат на его кодирование;

$D = \sum_{i,j,l} (w_{i,j}^{(l)} - \widehat{w}_{i,j}^{(l)})^2$ – квадратичная ошибка, а R –

битовые затраты, которые связаны с кодированием изображения.

Найденные в процессе RD-оптимизации значения q^* и \mathbf{k}^* далее используются при квантовании и арифметическом кодировании. Схема компрессии с использованием RD-оптимизации приведена на рис. 3.

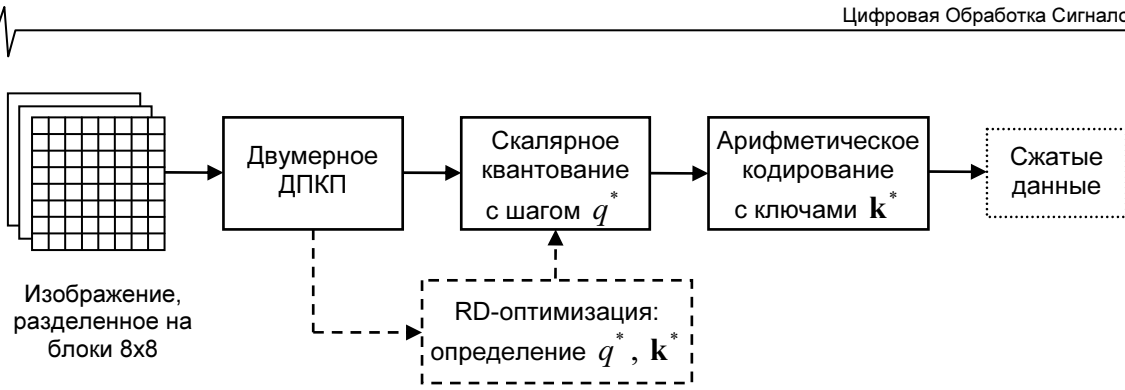


Рис. 3. Схема компрессии с RD-оптимизацией

Так как в алгоритме сжатия используется адаптивное арифметическое кодирование [6, 7], то для правильной оценки битовых затрат в процессе оптимизации требуется обновление внутренних гистограмм каждой модели без реального кодирования поступающих символов, то есть без непосредственного формирования выходного битового кода. Такую процедуру будем называть *виртуальным кодированием блока* (с заданным ключом $k_{i,j}$). Его алгоритм выглядит следующим образом:

1. Вычисление разности $d_{i,j}$ по формуле (4). Увеличение частоты символа $|d_{i,j}|$ в соответствующей модели.
2. Увеличение частоты символа $k_{i,j}$ в соответствующей модели.
3. Просматриваем последовательно все коэффициенты блока (за исключением постоянной составляющей), $l=1, \dots, 63$. Пусть $\tilde{w}_{i,j}^{(l)} \in Z^r$, тогда, если $k_{i,j}(r)=1$, увеличить частоту символа $|\tilde{w}_{i,j}^{(l)}|$ в модели, определяемой по правилу (6).

Тогда битовые затраты R_k на кодирование k -го символа алфавита можно оценить, используя соответствующую гистограмму (статистическую модель) арифметического кодера:

$$R_k = -\log_2 v_k,$$

где v_k – частота (то есть оценка вероятности) появления k -го символа алфавита.

Искать решение задачи (9) будем, проводя последовательную минимизацию:

$$L(\mathbf{k}^*, q^*) = \min_q \min_{\mathbf{k}} (L(\mathbf{k}, q) = D(\mathbf{k}, q) + \lambda R(\mathbf{k}, q)). \quad (10)$$

Пусть $L_{i,j}$ – значение функции Лагранжа для блока с координатами (i, j) , $D_{i,j}$ и $R_{i,j}$ – ошибка квантования и битовые затраты на кодирование блока с координатами (i, j) соответственно. Обозначим как $P_m(w)$ частоту появления символа w в модели с номером m , используемой для кодирования абсолютных значений коэффициентов ДПКП; $P_{dif}(u)$, $P_{key}(v)$ – частоты появления символов u и v в моделях кодирования абсолютных значений разностей ДИКМ постоянных составляющих и ключей блоков соответственно.

Поиск набора ключей блоков \mathbf{k}^* , минимизирующего функцию Лагранжа изображения

$$L = \sum_{i,j} L_{i,j} \text{ при заданных значениях } q \text{ и } \lambda,$$

выполняется по следующему алгоритму (последовательно для каждого блока):

1. Определить базовый ключ блока $k_{i,j}$ в соответствии с правилом (5). Положить текущее значение ключа

$k = k_{i,j}$, значение функции Лагранжа $L_{i,j}^* = +\infty$.

2. Вычислить функцию Лагранжа для блока при текущем значении ключа k :

$$L_{i,j} = D_{i,j} + \lambda R_{i,j},$$

$$D_{i,j} = \left(w_{i,j}^{(0)} - \tilde{w}_{i,j}^{(0)} \right)^2 + \sum_{l=1}^{63} \sum_{\tilde{w}_{i,j}^{(l)} \in Z_{i,j}^r, k(r)=1} \left(w_{i,j}^{(l)} - \tilde{w}_{i,j}^{(l)} \right)^2 +$$

$$+ \sum_{l=1}^{63} \sum_{\tilde{w}_{i,j}^{(l)} \in Z_{i,j}^r, k(r)=0} \left(w_{i,j}^{(l)} \right)^2$$

$$R_{i,j} = - \sum_{l=1}^{63} \sum_{\tilde{w}_{i,j}^{(l)} \in Z_{i,j}^r, k(r)=1} \log_2 P_m \left(|\tilde{w}_{i,j}^{(l)}| \right) -$$

$$- \log_2 P_{dif} \left(|d_{i,j}| \right) - \log_2 P_{key} (k) + a$$

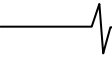
где $m = \text{Model}(\tilde{w}_{i,j}^{(l)})$ (6); a – количество ненулевых коэффициентов в блоке (считаем, что для кодирования знака каждого ненулевого коэффициента необходим 1 бит).

1. Если $L_{i,j} < L_{i,j}^*$, то положить $L_{i,j}^* = L_{i,j}$, $k_{i,j}^* = k$.
2. Выполнить полный перебор ключей $\{k\}$, которые могут быть получены из базового $k_{i,j}$ в результате всевозможных вариантов обнуления единичных битов базового ключа $k_{i,j}$: с каждым возможным текущим значением ключа k выполнить переход на шаг 2. Если перебор вариантов осуществлен, то перейти на шаг 5.

3. Выполнить виртуальное кодирование блока с ключом $k_{i,j}^*$ (по описанному выше алгоритму).

В результате работы алгоритма при заданных значениях параметра λ и шага квантования q будет получен набор оптимальных ключей блоков \mathbf{k}^* , в результате чего некоторая часть исходно ненулевых зон в блоках может быть впоследствии принудительно обнулена. Приведенную процедуру необходимо итерационно использовать при поиске оптимального значения шага квантования q^* , воспользовавшись одним из методов одномерной минимизации для функции $L(\mathbf{k}^*(q), q)$ по переменной q .

Количество итераций в описанной выше процедуре RD-оптимизации можно уменьшить за счет сокращения интервалов поиска оптимальных значений шага квантования и ключей. Так, например, эксперименты показали, что перебор всевозможных вариантов обнуления единичных битов на шаге 4 не требуется. Достаточно рассмотреть ключи, получаемые из базового заменой на каждой итерации одного бита, равного 1, на 0. То есть если число ненулевых битов в некотором ключе k_{ij} равно n , то вместо 2^n возможного коли-



чества вариантов обнуления зон достаточно рассмотреть n вариантов. Помимо этого, при реализации наиболее вычислительно емкого шага 2 процедуры RD-оптимизации возможно (и целесообразно) использование табличного вычисления квадратичной функции и функции логарифма.

Также было установлено, что связь между значением оптимального шага квантования q^* и параметра λ может быть приближенно описана соотношением: $q^* \approx C\sqrt{\lambda}$, где C – некоторая общая для различных изображений константа. При обработке изображений «Lena», «Barbara» и «Goldhill» с использованием метода наименьших квадратов было найдено следующее соотношение: $q^* \approx 2,21\sqrt{\lambda}$, поэтому поиск оптимального значения шага квантования q при минимизации $L(\mathbf{k}^*(q), q)$ выполняется в интервале $q \in (2\sqrt{\lambda}, 3\sqrt{\lambda})$.

Приведенный алгоритм RD-оптимизации вносит некоторую дополнительную вычислительную нагрузку в исходный алгоритм сжатия, но вместе с тем, существенно улучшает результаты базового алгоритма. Кроме того, усложнение алгоритма компрессии за счет введения процедуры RD-оптимизации относится только к стороне кодера, а объем операций, необходимых на стороне декодера, остается прежним.

Заключительный этап реализации схемы рис. 3 – много-модельное арифметическое кодирование каждого блока (i, j) с ключом $k_{i,j}^*$, найденным в результате RD-оптимизации – состоит из следующих шагов:

1. Арифметическое кодирование постоянной составляющей. Вычисление разности $d_{i,j}$ по формуле (4). Абсолютное значение разности $|d_{i,j}|$ кодируется в отдельном потоке. Если $d_{i,j} \neq 0$, то знак $s_{i,j} = \text{sign}(d_{i,j})$ (1 бит) записывается в поток знаков.

2. Арифметическое кодирование ключа $k_{i,j}^*$ в соответствующем потоке.

3. Арифметическое кодирование спектральных коэффициентов из ненулевых зон. Просматриваем последовательно все коэффициенты блока (за исключением постоянной составляющей), $l=1, \dots, 63$. Пусть $\hat{w}_{i,j}^{(l)} \in Z_{i,j}^r$, тогда, если $k_{i,j}^*(r) = 1$, закодировать $|\hat{w}_{i,j}^{(l)}|$ арифметическим кодером, используя модель, определяемую по правилу (6), и, если $\hat{w}_{i,j}^{(l)} \neq 0$, записать $s_{i,j}^{(l)} = \text{sign}(\hat{w}_{i,j}^{(l)})$ в файл знаков. Если $k_{i,j}^*(r) = 0$, то коэффициент $\hat{w}_{i,j}^{(l)}$ не обрабатывается (и будет восстановлен декодером как $\hat{w}_{i,j}^{(l)} = 0$).

Экспериментальные результаты и выводы

На графиках (рис. 4, 5 и 6) приведены результаты, полученные при сжатии изображений «Goldhill», «Barbara» и «Lena» базовым и модифицированным вариантами алгоритма, а также по методу JPEG. Сжатие по стандарту JPEG выполнялось с помощью программного обеспечения от Independent JPEG Group (версия 6b) [16] с использованием библиотеки арифметического кодирования [17]. Для корректности сравнения при расчете битовых затрат размер заголовка вычитался из общего размера файла JPEG.

Ошибка кодирования изображений (ось ординат) на рис. 4-6 измерялась по величине пикового отношения сигнала к шуму:

$$\text{PSNR} = 20 \lg \frac{255}{\sqrt{\frac{1}{MN} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} (x_{ij} - \tilde{x}_{ij})^2}} \quad (\text{дБ}),$$

где $\mathbf{X} = \{x_{ij}\}$ – исходное, $\tilde{\mathbf{X}} = \{\tilde{x}_{ij}\}$ – восстановленное изображение размера $M \times N$ пикселей. Величина сжатия данных (ось абсцисс) выражена в битовых затратах на кодирование одного пикселя изображения, bpp (bits per pixel).

Как показывают приведенные результаты для распространенных тестовых изображений, предложенный алгоритм компрессии изображений по эффективности не уступает JPEG на базе ДКП, более того, на мелкодетальных «высокочастотных» изображениях (типа «Barbara») показывает преимущества в сжатии при меньших вычислительных затратах, а следовательно, в приложениях, имеющих ограничения на аппаратные ресурсы, может заменить метод JPEG.

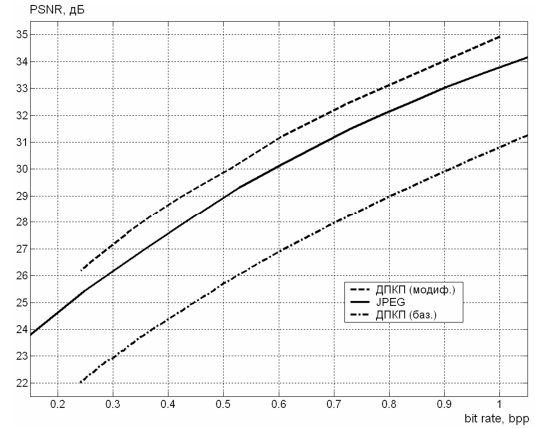


Рис. 4. График зависимости величины PSNR (дБ) от количества бит на пиксель изображения (bpp) («Barbara»)

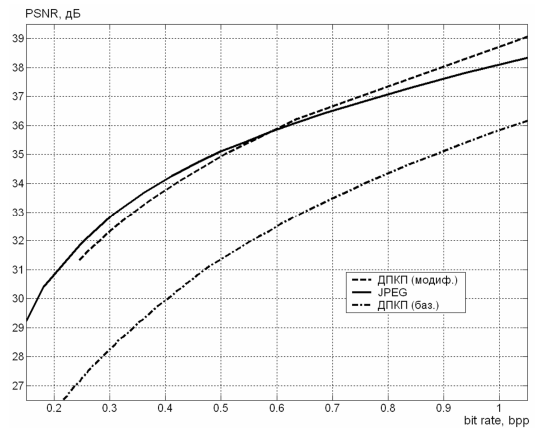


Рис. 5. График зависимости величины PSNR (дБ) от количества бит на пиксель изображения (bpp) («Lena»)

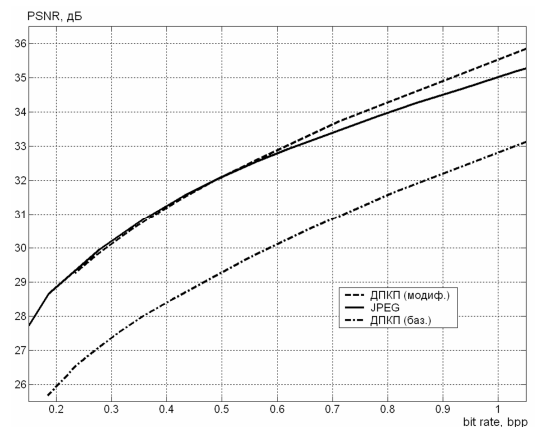


Рис. 6. График зависимости величины PSNR (дБ) от количества бит на пиксель изображения (bpp) («Goldhill»)

Литература

1. E. Feig, S. Winograd. On the multiplicative complexity of discrete cosine transform // IEEE transactions on information theory. – July 1992. – Vol. 38. – No. 4. – p. 1387-1391.
2. Jie Liang, Trac D. Tran. Fast multiplierless approximations of the DCT with the lifting scheme // IEEE transactions on signal processing. – December 2001. – Vol. 49. – No. 12. – p. 3032 – 3044.
3. Neri Merhav. Multiplication-free approximate algorithms for compressed domain linear operations on images // IEEE transactions on image processing. – February 1999. – Vol. 8. – No. 2. – p. 247-254.
4. С. В. Умняшкин. О модификации дискретного косинусного преобразования // Изв. Тул. гос. ун-та. Сер. Математика. Механика. Информатика. Тула: ТулГУ. – 1998. – Т. 4. – Вып. 1. – с. 143-147.
5. N. Jayant, P. Noll. Digital coding of waveforms. – Englewood Cliffs, NJ: Prentice Hall, 1984.
6. I. H. Witten, R. M. Neal, J. G. Cleary. Arithmetic coding for data compression // Communications of the ACM. – June 1987. – Vol. 30. – No. 6. – p. 520-540.
7. M. Nelson, J.-I. Gailly. The data compression book (second edition). – New York: M&T Books, 1995. – p. 541.
8. J. Rissanen. K. M. Mohiuddin. A multiplication-free multialphabet arithmetic code // IEEE transactions on communication. – February 1989. – Vol. 37. – p. 93-98.
9. М. В. Захаров. Об одной реализации арифметического кода. Электронный ресурс <http://www.maxime.net.ru/doc/mfarc.ps.gz>.
10. G. Davis, A. Nosratinia. Wavelet-based image coding: an overview // Applied and computational control, signals and circuits. – 1998. – Vol. 1. – No. 1. – p. 205-269.
11. С. В. Умняшкин. Использование контекстного арифметического кодирования для повышения сжатия данных по схеме JPEG // Известия вузов. Электроника. – 2001. – №3. – с. 96-99.
12. Ф. В. Стрелков, С. В. Умняшкин. Контекстное кодирование коэффициентов дискретного косинусного преобразования в JPEG-подобной схеме компрессии // Цифровая обработка сигналов. – 2003. – № 2. – с. 5-10.
13. X. Xue, W. Gao. Context-based statistical model for DCT-based image coder // Proceedings of picture coding symposium '99. Oregon State Univ. Corvallis, OR, USA. – April 1999. – p. 383-385.
14. С. В. Умняшкин, Е. А. Безуглова. Контекстное кодирование коэффициентов дискретного косинусного преобразования на основе межблочной корреляции в JPEG-подобной схеме компрессии // Цифровая обработка сигналов. – 2004. – № 2. – с. 13-17.
15. T. Berger. Rate distortion theory. – Englewood Cliffs, NJ: Prentice Hall, 1971.
16. Электронный ресурс <http://www.ijg.org>.
17. Электронный ресурс <http://sylvania.net/jpeg-ari/jpeg-ari.zip>.

A NEW STILL IMAGE COMPRESSION ALGORITHM BASED ON DCT MODIFICATION**Sergei V. Umnyashkin, Valentina V. Kurina.**

A new still image compression algorithm based on fast integer multiplication free DCT modification is proposed in this paper. The main features of the algorithm are RD-optimization, which is used to define optimal parameters of the algorithm, and arithmetic coding with several models. Developed algorithm is not inferior in performance to JPEG based on DCT and demands less computational costs.

Уважаемые коллеги!

Вы можете подписаться на наш журнал на первое полугодие 2010 года в любом отделении почтовой связи. Подписной индекс издания по каталогу ОАО «Роспечать» - 82185. Журнал издается ежеквартально.

Для тех, кто не успел оформить подписку на первое полугодие 2010 года через ОАО «Роспечать», сохраняется возможность приобретения журналов непосредственно в редакции по адресу: 107031, г. Москва, Рождественка, 6\9\20, стр. 1, Российское научно-техническое общество радиотехники, электроники и связи им. А.С. Попова, или оформить Заказ в соответствии с требованиями, выставленными на сайте журнала: www.dsra.ru.

Справки по телефонам: (495) 621-71-08, 621-06-10.

Факс: (495) 621-16-39.

E-mail: nto.popov@mtu-net.ru

tor@rgta.ryazan.ru

СТЕГАНОГРАФИЧЕСКОЕ СКРЫТИЕ ИНФОРМАЦИИ В СТАТИЧЕСКИХ ИЗОБРАЖЕНИЯХ

Богущ Р.П., к.т.н., доцент, зав. кафедрой вычислительных систем и сетей Полоцкого государственного университета, Беларусь, bogushr@mail.ru

Ключевые слова: стеганографическая система, статические изображения, дискретное преобразование, скрытие информации.

Введение

Широкое распространение мультимедийных технологий, сети Internet, непрерывное совершенствование компьютерной техники, методов обработки цифровой информации определило развитие стеганографии, которая позволяет обеспечивать обмен конфиденциальной информацией таким образом, что скрывается сам факт передачи такой информации.

В цифровой стеганографии используются следующие термины [1,2]:

- стегосистема – система скрытия данных на основе стеганографии;
- цифровой контейнер - любая цифровая информация, предназначенная для сокрытия конфиденциальных сообщений;
- скрытое сообщение – сообщение, встраиваемое в контейнер;
- стегоконтейнер - контейнер с внедренным в него скрытым сообщением.

Следует отметить, что используемый цифровой контейнер оказывает значительное влияние на безопасность передаваемой информации. Среди ряда возможных стегоконтейнеров - статические и динамические изображения. Это обусловлено наличием в большинстве реальных изображений текстурных областей, имеющих шумовую структуру и, следовательно, пригодных для встраивания дополнительной информации, слабой чувствительностью человеческого глаза к незначительным изменениям яркостных, цветовых и др. характеристик изображения, а также бурно развивающимися методами цифровой обработки изображений.

Стегосистема, использующая в качестве цифрового контейнера статическое изображение, должна удовлетворять следующим основным требованиям: свойства контейнера должны быть модифицированы так, чтобы изменения невозможно было определить при визуальной атаке; стегоконтейнер должен быть устойчив преднамеренным и непреднамеренным искажениям - сжатию с потерей информации, преобразованию в другой формат и т. д. В настоящее время существует ряд стеганографических систем [1-5], однако совершенствование методов обработки и сжатия изображений требует непрерывного их развития, так как чем более совершенными становятся методы сжатия, тем меньше остается возможностей для внедрения дополнительной инфор-

Рассматривается стеганографическая система скрытия информации, представленной в виде изображения в цветных статических изображениях. Для скрытия информации используется дискретное косинусное преобразование либо двумерное вейвлет-преобразование, адаптированные к современным алгоритмам сжатия изображений JPEG и JPEG2000 соответственно. Криптостойкость системы обеспечивается применением ассиметричного алгоритма RSA. Рассматриваются вопросы эффективности и робастности предлагаемой системы. Представлены результаты исследований, которые подтвердили возможность ее использования в системах скрытой передачи информации.

мации. Поэтому в данном направлении непрерывно проводятся исследования, а число публикаций по данной тематике постоянно возрастает. Наиболее эффективны стегосистемы, реализующие скрытие данных в области спектральных преобразования с учетом особенностей алгоритмов сжатия изображений [3-5].

Целью данной работы является разработка и исследование алгоритмов скрытия данных в изображениях на основе дискретного косинусного преобразования и вейвлет-преобразования для стандартов сжатия JPEG и JPEG2000.

Внедрение и восстановление скрываемой информации

В разработанной стегосистеме сообщение, представляемое в виде изображения, внедряется в цифровой контейнер (статическое изображение) путем изменения амплитуд спектральных коэффициентов, полученных в результате прямого дискретного преобразования на основе выражения:

$$E = S + \frac{A}{255} \cdot (W - 128), \quad (1)$$

где S - амплитуда спектральной составляющей цифрового контейнера, E - амплитуда спектральной компоненты стегоконтейнера, W - компонента скрываемого сообщения, A - весовой коэффициент, определяемый величиной энергии, которую можно добавить к изображению без его существенного искажения.

В связи с тем, что свойства контейнера должны быть модифицированы так, чтобы изменение невозможно было выявить при визуальном контроле, энергия, добавленная к изображению, должна рассчитываться на этапе внедрения информации. Определение весового коэффициента для контейнера размером $N \times M$ выполняется на основе выражения:

$$A = \sqrt{\frac{4 \cdot k}{N_1} \cdot \sum_{i=1}^{N \cdot M} |S_i|^2}, \quad (2)$$

где N_1 - общее число компонент скрываемого сообщения; k - коэффициент внедрения, подбираемый для контейнера и скрываемого сообщения.

Для обеспечения высокой криптостойкости стегосистемы можно использовать ассиметричный современный криптографический алгоритм RSA, безопасность которого основана на сложности разложения на множители больших чисел [8]. Для скрытия в стегоконтейнере встроенного сообщения с помощью алгоритма RSA и обеспечения высокого быстродействия шифруются номера спектральных составляющих, в которых содержатся компоненты скрытого сообщения. При этом перед шифрованием каждый номер соответствующей спектральной составляющей дополняется строкой случайных бит длиной M (в зависимости от длины выбранного модуля).

Известно, что реальные изображения не являются случайным процессом с равномерно распределенными значениями величин, и большая часть энергии изображений сосредоточена в низкочастотной части спектра. Высокочастотные составляющие наиболее подвержены воздействию со стороны различных алгоритмов обработки, будь то сжатие, низкочастотная (размытие) или медианная фильтрация. Поэтому, для обеспечения устойчивости стегосистемы к алгоритмам сжатия использованы низкочастотные компоненты контейнера.

Таким образом, алгоритм скрытия сообщения требует выполнения следующих основных шагов: выполнение прямого дискретного преобразования для контейнера - ДКП при использовании алгоритма JPEG и вейвлет-преобразования при использовании JPEG2000; расчёт весового коэффициента для контейнера с использованием выражения (2); скрытие сообщения с использованием выражения (1) и алгоритма RSA; выполнение обратного дискретного преобразования.

Восстановление скрываемого сообщения требует наличие контейнера, стегоконтейнера и закрытого ключа алгоритма RSA. При этом требуется выполнить прямое дискретное преобразование для стегоконтейнера и извлечь сообщение с использованием закрытого ключа алгоритма RSA на основе выражения:

$$W = \frac{255}{A} \cdot (E - S) + 128,$$

где S - амплитуда спектральной составляющей цифрового контейнера, E - амплитуда спектральной компоненты стегоконтейнера,

Экспериментальные исследования

На первом этапе экспериментов при скрытии информации использовалось дискретное косинусное преобразование, а стегоконтейнеры подвергались JPEG сжатию и преобразованию в формат GIF и обратно. По результатам исследований установлено, что JPEG сжатие приводит к более существенным искажениям встроенных сообщений по сравнению с преобразованием палитры стегоконтейнера (примеры показаны на рис. 1 - 3).

На втором этапе экспериментов при скрытии информации использовалось вейвлет-преобразование, а стегоконтейнеры, полученные с различными весовыми коэффициентами, подвергались JPEG2000 кодированию с разными степенями сжатия. По результатам исследований установлено, что для используемого стегоконтейнера необходим выбор оптимального весового коэффициента на основе визуального контроля стегоконтейнера при внедрении, а также установлено, что значительное увеличение степени сжатия все же приводит к потерям качества скрываемой информации. Поэтому дальнейшее совершенствование разработанной стегосистемы предполагается за счет применения помехоустойчивого кодирования. Результаты экспериментов представлены на рис. 4 и в таблицах 1 - 2. Размер цветного контейнера (24 бита на пиксель) составляет 305×200 пикселей. Внедряемое сообщение (рис. 4.1в) размером 42×39 пикселей (на рис. 4.1в и далее представлено увеличенным в два раза) использовалось без рамки, она показана на рисунке для наглядности границ. В ходе экспериментов рассчитывалась степень сходства восстановленного сообщения $D' = \{d'_{ij}\}$ и внедряемого $D = \{d_{ij}\}$ по формуле:

$$R^{SSD} = 1 - \frac{1}{256} \frac{\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} (d_{ij} - d'_{ij})^2}{\sqrt{\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} (d_{ij})^2} \sqrt{\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} (d'_{ij})^2}}.$$



Рис.1. Контейнер



Рис.2. Стегоконтейнер



а)



б)



в)

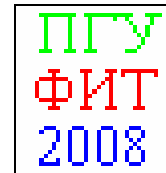
Рис.3. Скрываемое изображение: а) исходный вид; б) восстановленное после JPEG сжатия стегоконтейнера; в) восстановленное после преобразования цветовой палитры стегоконтейнера.



а)



б)



в)

Рис. 4. Результаты экспериментов: а) цифровой контейнер (179КБ); б) стегоконтейнер (179КБ); в) скрытое сообщение (5 КБ).

Таблица 1

Размер файла стегоконтейнера, КБ	115	45	21	12	8
Весовой коэффициент					
50	0,996	0,995	0,989	0,958	0,899
100	0,996	0,995	0,989	0,965	0,916
200	0,996	0,995	0,989	0,973	0,942

Рассчитанные значения коэффициента схожести R^{SSD} для скрываемой и восстановленной информации приведены в Таблице 1, а в Таблице 2 даны восстановленные сообщения при различных степенях сжатия в формате JPEG2000.

Визуальный анализ результатов (таблица 2) и анализ рассчитанных коэффициентов схожести (таблица 1) показывает, что при оптимальном значении весового коэффициента для заданного контейнера значительное увеличение степени сжатия приводит к потерям качества восстановленного сообщения, в тоже время с увеличением весового коэффициента ухудшается качество стегоизображения.

Далее тестировалась робастность стегосистемы к потере или преднамеренной порче части стегоконтейнера при передаче или хранении. Стабильность стегосистемы проверялась следующим образом: в стегоконтейнере имитировалась потеря фрагмента изображения (до 30%) в произвольном месте. Результаты экспериментов показаны на рисунке 5 и в Таблице 3.

Из Таблицы 3 следует, что результаты экспериментов по робастности алгоритма к потере части информации стегоконтейнера.

Анализ полученных результатов показывает, что восстановленное изображение визуально распознаваемо даже при потере 25% контента стегоконтейнера.

Заключение

Для обеспечения робастности стегосистемы к компрессии изображений в форматах сжатия JPEG и JPEG-2000 предусмотрено возможность использования ДКП или вейвлет-преобразования на этапе внедрения скрываемого сообщения. Разработанный алгоритм скрытия информации использует низкочастотные компоненты изображения, что увеличивает его устойчивость к алгоритмам сжатия, при этом различие между стегоконтейнером и контейнером при оптимальном выборе параметров стегосистемы визуально установить достаточно сложно. Криптостойкость системы обеспечивается применением асимметричного криптографического алгоритма RSA при внедрении информации в контейнер. Восстановление скрываемого сообщения требует наличие контейнера, стегоконтейнера и закрытого ключа алгоритма RSA.

Проведен ряд экспериментов, которые подтвердили эффективность данной стегосистемы и показали ее перспективность для скрытой передачи информации, в системах встраивания цифровых водяных знаков и идентификационных номеров. Однако по результатам исследований установлено, что при оптимальном коэффициенте внедрения для заданного контейнера значительное увеличение степени сжатия или фильтрация стегоконтейнера приводит к потерям качества восстановленного сообщения. В связи с этим, дальнейшее совершенствование разработанной стегосистемы возможно путем применения помехоустойчивых кодов на этапе внедрения информации в контейнер.

Таблица 2

Размер файла стегоконтейнера, КБ Весовой коэффициент	115	45	21	12	8
50					
100					
200					

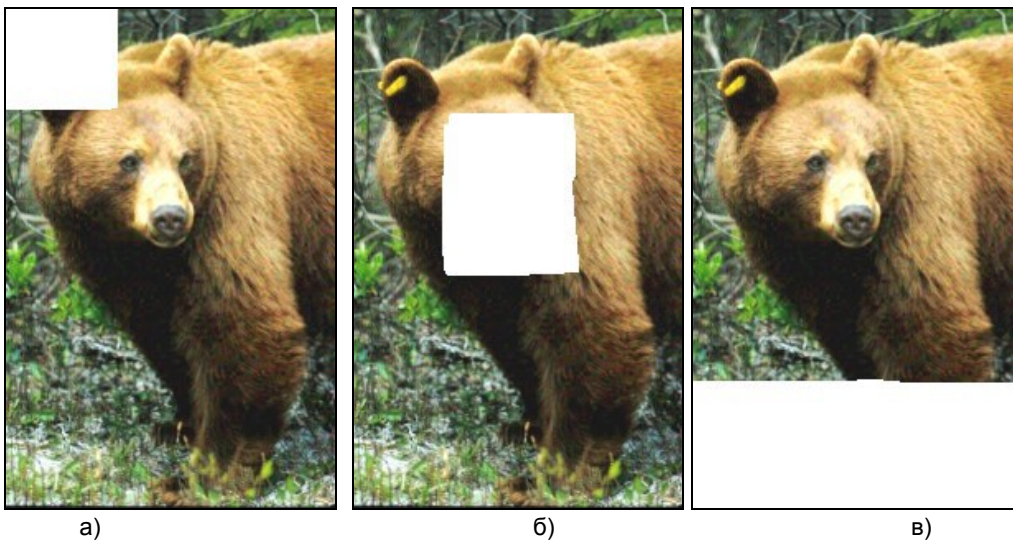


Рис.5. Стегоконтейнер с потерей части контента: а) с потерей 6,5 % контента в левом верхнем углу; б) с потерей 12,8 % контента в центральной части; в) с потерей 26 % контента в нижней части.

Таблица 3

Местоположение и количество утерянного изображения стегоконтейнера, %	в левом верхнем углу, 6,5 %	в центральной части, 12,8 %	в нижней части, 3 %	в нижней части, 13 %	в нижней части, 26 %
Восстановленное сообщение					
R^{SSD}	0,982	0,972	0,994	0,985	0,961

Литература

1. Грибунин, В.Г. Цифровая стеганография / В.Г.Грибунин, И.Н Оков, И.В. Туринцев. - М.: СОЛОН-Пресс, 2002. - 261с.
2. Кашеев, А.А. Стеганографическая защита цифровых изображений / А.А. Кашеев, С.Б. Саломатин // Известия Белорусской инженерной академии. - 2003. - №1(15)/1. - С.215 - 217.
3. Akansu, A.N. Data hiding in multimedia - theory and applications/ A.N. Akansu //Doctoral Dissertation Department of ECE New Jersey Institute of Technology University Heights, Newark, NJ 07032, 1999.
4. Pereira, S. A framework for optimal adaptive DCT watermarks using linear programming / S. Pereira, T. Pun // Proc. of Tenth European Signal Processing Conf., EUSIPCO'2000, Tampere, Finland, Sep. 5 - 8 2000.- Tampere,2000. - P.42 - 46
5. Hassanien,A. Watermarking for copyright protection using discrete wavelet transform / A.Hassanien // Proc. of the 8 Int. Conf. Pattern Recognition and Information Processing, PRIP'2005, Minsk, May 18-20 2005 / Belarusian State University of Informatics and Radioelectronics.-Minsk, 2005 - P.185-191
6. Ватолин, Д. Методы сжатия данных. Устройство архиваторов, сжатие изображений и видео / Д. Ватолин, А. Ратушняк, М. Смирнов, В. Юкин. - М.: ДИАЛОГ-МИФИ, 2002. - 384 с.
7. Chen, G. Applications of Wavelet Transforms in Pattern Recognition and De-noising / G. Chen.- Montreal, Concordia University, Canada, 1999. - 120p.
8. Шнайер, Б. Прикладная криптография. Протоколы, алгоритмы, исходные тексты на языке Си / Б. Шнайер.- М.: Триумф, 2002. - 816 с.

МЕТОД ВЫДЕЛЕНИЯ КОНТУРОВ В ИЗОБРАЖЕНИИ НА ОСНОВЕ ВЫЧИСЛЕНИЯ КОЛИЧЕСТВА ИНФОРМАЦИИ

Медведева Е. В., к.т.н., доцент кафедры радиоэлектронных средств Вятского государственного университета, EMedv@mail.ru

Ключевые слова: количество информации, контур, изображение, коррекция, точечная помеха, метод выделения.

Введение

Для распознавания объекта на изображении надо искать набор частей изображений, которые соответствуют частям объекта и удовлетворяют соответствующим ограничениям. Один из возможных методов определения частей объектов основан на выделении контуров или границ на изображении [1,2], поскольку по контурам можно определить размер, форму, положение объекта, сравнить данные с последующими или предыдущими изменениями и т.д. Кроме того, так как изображение в виде контуров не содержит полутонов и мелких деталей, то при сжатии известными алгоритмами компрессии для хранения и передачи требуется меньший объем бит, чем для цифрового полутонового изображения (ЦПИ).

В данной работе предложен метод выделения контуров в ЦПИ, представленных g -разрядными двоичными числами, основанный на вычислении величины количества информации в элементах двоичного изображения и имеющий низкую вычислительную сложность.

Метод выделения контуров

Для вычисления количества информации в элементах изображения в качестве математической модели ЦПИ выбрано одностороннее марковское случайное поле (ОМСП), называемое также двумерной марковской цепью на несимметричной полуплоскости [3,4].

Представление ЦПИ набором из g -разрядных двоичных изображений (РДИ) позволяет свести задачу построения математической модели ЦПИ к построению математической модели g -РДИ, каждое из которых представляет собой однородную двумерную цепь Маркова с двумя равновероятными ($p_1^{(i)} = p_2^{(i)}$) значениями $M_1^{(i)}$ и $M_2^{(i)}$.

Будем считать, что l -е РДИ ($l = \overline{1, g}$) представляет собой случайное марковское поле с разделимой автокорреляционной функцией вида [3]

$$r_{i,j} = \sigma_\mu^2 \exp\{-\alpha_1 |f| - \alpha_2 |s|\}, \quad (1)$$

где i, j – дискретные пространственные координаты (по горизонтали и вертикали); σ_μ^2 – дисперсия двумерного дискретнозначного марковского процесса; α_1, α_2 – множители, зависящие от ширины спектральной плотности мощности случайных процессов по двум измерениям; f, s – шаг корреляции по горизонтали и вертикали.

В этом случае РДИ можно представить как суперпозицию двух одномерных цепей Маркова по горизонтали и вертикали с двумя равновероятными значениями $M_1^{(i)}$, $M_2^{(i)}$ и матрицами вероятностей переходов от значения $M_i^{(i)}$ к соседнему значению $M_j^{(i)}$ по горизонтали и верти-

Предложен метод выделения контуров в цифровых полутоновых изображениях, представленных g -разрядными двоичными числами, основанный на вычислении количества информации в элементах двоичного изображения. Для коррекции контура изображения предложен метод удаления ложных контуров – точечных помех. Приведены результаты моделирования.

кали изображения, соответственно:

$${}^1\Pi^{(l)} = \begin{vmatrix} {}^1\pi_{11}^{(l)} & {}^1\pi_{12}^{(l)} \\ {}^1\pi_{21}^{(l)} & {}^1\pi_{22}^{(l)} \end{vmatrix}, \quad {}^2\Pi^{(l)} = \begin{vmatrix} {}^2\pi_{11}^{(l)} & {}^2\pi_{12}^{(l)} \\ {}^2\pi_{21}^{(l)} & {}^2\pi_{22}^{(l)} \end{vmatrix}. \quad (2)$$

На рис. 1 представлено ОМСП ЦПИ, разделенное на области $F_i^{(l)}$ ($i = \overline{1, 4}$), элементы которых являются цепями Маркова различной размерности. Вычисление количества информации в элементах изображения областей $F_1^{(l)}$, $F_2^{(l)}$ и $F_3^{(l)}$ хорошо известно [3,4]. Наибольшую сложность представляет алгоритм вычисления количества информации элементов области $F_4^{(l)}$, так как элемент $\mu_{i,j}^{(l)}$ области $F_4^{(l)}$ зависит от трех соседних элементов, входящих в окрестность $\Lambda_{i,j}^{(l)} = \{\mu_{i,j-1}^{(l)}, \mu_{i-1,j}^{(l)}, \mu_{i-1,j-1}^{(l)}\}$.

На рис. 2 приведен фрагмент двумерного бинарного изображения, соответствующего области $F_4^{(l)}$ ОМСП (рис. 1).

Для вычисления количества информации в элементе $\nu_4^{(l)}$ относительно элементов $\nu_3^{(l)}, \nu_2^{(l)}, \nu_1^{(l)}$ l -го РДИ (рис.2) в работах [3,4] получено выражение вида:

$$I(\nu_4^{(l)} | \nu_3^{(l)}, \nu_2^{(l)}, \nu_1^{(l)}) = -\log \frac{{}^1\pi_{ii}^{(l)} {}^2\pi_{ii}^{(l)}}{{}^3\pi_{ii}^{(l)}}, \quad (3)$$

где ${}^3\pi_{ii}^{(l)}$; $i, j = \overline{1, 2}; i \neq j$ элементы матрицы ${}^3\Pi = {}^1\Pi \cdot {}^2\Pi = \begin{vmatrix} {}^3\pi_{ii}^{(l)} \\ {}^3\pi_{ij}^{(l)} \end{vmatrix}$.

Для элементов $l^{2 \times 2}$ -го РДИ, принадлежащих области $F_4^{(l)}$ (рис.1), количество информации между элементом $\nu_4^{(l)}$ и различными сочетаниями значений элементов окрестности $\Lambda_{ij}^{(l)}$ можно определить по матрице Π (4):

$$\Pi = \begin{vmatrix} \alpha_1 & \alpha'_1 \\ \alpha_2 & \alpha'_2 \\ \alpha_3 & \alpha'_3 \\ \alpha_4 & \alpha'_4 \\ \beta_1 & \beta'_1 \\ \beta_2 & \beta'_2 \\ \beta_3 & \beta'_3 \\ \beta_4 & \beta'_4 \end{vmatrix} = \begin{vmatrix} \frac{{}^1\pi_{ii}^{(l)} {}^2\pi_{ii}^{(l)}}{{}^3\pi_{ii}^{(l)}} & \frac{{}^1\pi_{ij}^{(l)} {}^2\pi_{ij}^{(l)}}{{}^3\pi_{ii}^{(l)}} \\ \vdots & \vdots \\ \frac{{}^1\pi_{ij}^{(l)} {}^2\pi_{ij}^{(l)}}{{}^3\pi_{ii}^{(l)}} & \frac{{}^1\pi_{ii}^{(l)} {}^2\pi_{ii}^{(l)}}{{}^3\pi_{ii}^{(l)}} \\ \frac{{}^1\pi_{ii}^{(l)} {}^2\pi_{ii}^{(l)}}{{}^3\pi_{ii}^{(l)}} & \frac{{}^1\pi_{ij}^{(l)} {}^2\pi_{ij}^{(l)}}{{}^3\pi_{ii}^{(l)}} \\ \frac{{}^1\pi_{ii}^{(l)} {}^2\pi_{ii}^{(l)}}{{}^3\pi_{ii}^{(l)}} & \frac{{}^1\pi_{ij}^{(l)} {}^2\pi_{ij}^{(l)}}{{}^3\pi_{ii}^{(l)}} \\ \vdots & \vdots \\ \frac{{}^1\pi_{ij}^{(l)} {}^2\pi_{ij}^{(l)}}{{}^3\pi_{ii}^{(l)}} & \frac{{}^1\pi_{ii}^{(l)} {}^2\pi_{ii}^{(l)}}{{}^3\pi_{ii}^{(l)}} \end{vmatrix}, \quad \begin{matrix} \nu_1\nu_2\nu_3 \rightarrow \nu_4 \\ 000 \rightarrow 0 \\ \vdots \\ 111 \rightarrow 0 \\ 111 \rightarrow 1 \\ \vdots \\ 000 \rightarrow 1 \end{matrix} \quad (4)$$

где $\alpha'_i = 1 - \alpha_i$, $\alpha_i = \beta_i$ ($i = \overline{1, 4}$), $\alpha_4 = \alpha'_1$ и $\beta_4 = \beta'_1$.

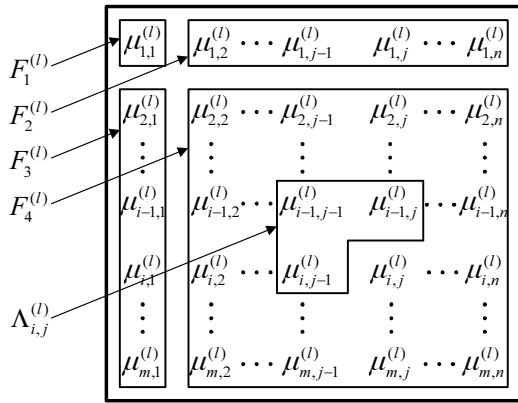
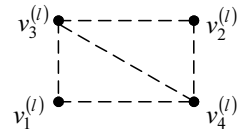


Рис. 1. Области ОМСП с окрестностью из трех элементов



$$\begin{aligned} v_1^{(l)} &= \mu_{i,j-1}^{(l)}; \\ v_2^{(l)} &= \mu_{i-1,j}^{(l)}; \\ v_3^{(l)} &= \mu_{i-1,j-1}^{(l)}; \\ v_4^{(l)} &= \mu_{i,j}^{(l)}. \end{aligned}$$

Рис. 2. Фрагмент двумерного бинарного изображения

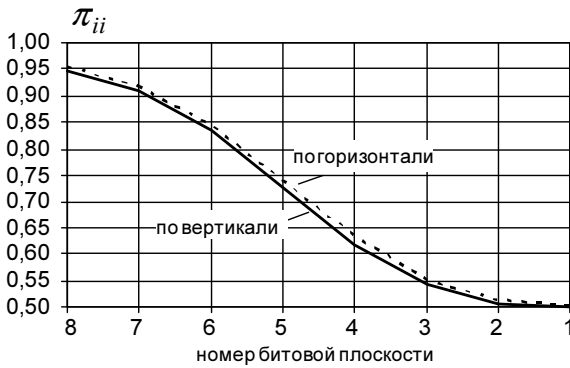


Рис.3. Усредненное распределение вероятностей переходов (8 – старший разряд; 1 – младший разряд)

Для вычисления количества информации в элементах двоичного изображения необходимо знать элементы матриц вероятностей переходов по горизонтали и вертикали (2). На рис. 3 приведено усредненное распределение элементов матриц вероятностей переходов в g РДИ на основе большого числа выборов реальных изображений.

Очевидно, что количество информации в элементе РДИ будет минимально, если окрестные элементы V_1, V_2 имеют знаки одинаковые с V_4 [4,5].

В случае появления на РДИ областей другой яркости, на границе области один или два окрестных элемента будут иметь разные с V_4 знаки, и количество информации в элементе v_4 увеличивается. Сравнивая значения вычисленной величины количества информации в элементе изображения с порогом, определяем, будет ли данная точка являться точкой контура.

Значение порога H вычисляют для каждого РДИ с учетом вычисленного минимального количества информации и количества информации, когда хотя бы один из элементов окрестности будет иметь другие знаки:

$$I(v_4 = M_1 | v_1 = M_1; v_2 = M_1) < H < I(v_4 = M_1 | v_1 = M_1; v_2 = M_2). \quad (5)$$

Например, для старшего (восьмого) РДИ при ${}^1\pi_{ii}^{(l)} = {}^2\pi_{ii}^{(l)} = 0,95$, если окрестные элементы V_1 и V_2 имеют знаки одинаковые с V_4 , то количество информации в элементе изображения будет минимально:

$$I(v_4 | v_1, v_2) = -\log \left(\frac{{}^1\pi_{ii}^{(l)} {}^2\pi_{ii}^{(l)}}{{}^3\pi_{ii}^{(l)}} \right) = -\log \left(\frac{0,95^2}{0,905} \right) = 0,004 \text{ бит.}$$

При появлении границы на изображении, в случае, если только один из окрестных элементов совпадает по знаку с элементом V_4 , то количество информации в элементе изображения равно:

$$I(v_4 | v_1, v_2) = -\log \left(\frac{{}^1\pi_{ii}^{(l)} {}^2\pi_{ij}^{(l)}}{{}^3\pi_{ij}^{(l)}} \right) = -\log \left(\frac{0,95 \cdot 0,05}{0,095} \right) = 1 \text{ бит.}$$

А значение порога H определится из условия (5): $0,004 < H < 1$.

Толщина линии контура на одном РДИ будет составлять 1 элемент. Известно, что основные детали области можно выделить на старших РДИ (при $\pi_{ii} \rightarrow 1$), а младшие РДИ (при $\pi_{ii} \rightarrow 0,5$) будут составлять фон изображения в виде двумерного шума. На основе моделирования большой совокупности ЦПИ было определено, что для определения контуров изображений достаточно двух старших РДИ (8-го и 7-го).

На рис.4 приведен пример определения контуров на ЦПИ.



а)



б)

Рис. 4. Исходное ЦПИ (а) и изображение с выделенными контурами (б)

Метод удаления ложных контуров

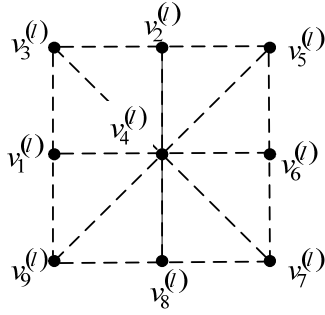


Рис. 5. Фрагмент двумерного бинарного изображения

В результате выделения контуров появляются ложные контура – точечные помехи. С целью улучшения контура сначала необходимо осуществить коррекцию изображения – удалить точечные помехи, а затем выделить контуры контрастных областей. Для удаления точечных помех предлагается вычислить количество информации в элементе v_4 относительно элементов окрестности

$$\Lambda_{ij}^* = \{v_1^{(l)}, v_2^{(l)}, v_3^{(l)}, v_5^{(l)}, v_6^{(l)}, v_7^{(l)}, v_8^{(l)}, v_9^{(l)}\} \text{ (рис. 5).}$$

По аналогии с выражением (3), получена формула для вычисления количества информации в элементе v_4 относительно элементов окрестности Λ_{ij}^* .

$$I(v_4^{(l)} | v_1^{(l)}, v_2^{(l)}, v_6^{(l)}, v_8^{(l)}) = -\log \left(\frac{{}^1\pi_{ii}^{(l)2} {}^2\pi_{ii}^{(l)6} {}^6\pi_{ii}^{(l)8} {}^8\pi_{ii}^{(l)10} {}^{\overline{10}}\pi_{ii}^{(l)}}{{}^3\pi_{ii}^{(l)5} {}^5\pi_{ii}^{(l)7} {}^7\pi_{ii}^{(l)9} {}^{\overline{9}}\pi_{ii}^{(l)}} \right), \quad (6)$$

где ${}^q\pi_{ij}^{(l)}$ ($i, j = 1, 2; i \neq j; q = \overline{1, 10}$) – элементы матриц вероятностей переходов ${}^1\Pi$, ${}^6\Pi$ – по горизонтали; ${}^2\Pi$, ${}^8\Pi$ – по вертикали;

$${}^3\Pi^{(l)} = {}^1\Pi^{(l)} \cdot {}^2\Pi^{(l)}; \quad {}^5\Pi = {}^2\Pi \cdot {}^6\Pi; \quad {}^7\Pi = {}^6\Pi \cdot {}^8\Pi;$$

$${}^9\Pi = {}^1\Pi \cdot {}^8\Pi;$$

$${}^{10}\Pi = {}^1\Pi \cdot {}^2\Pi \cdot {}^6\Pi \cdot {}^8\Pi = {}^3\Pi \cdot {}^7\Pi = {}^5\Pi \cdot {}^9\Pi.$$

Для элементов l -го РДИ количество информации между элементом $v_4^{(l)}$ и различными сочетаниями значений элементов окрестности Λ_{ij}^* можно определить по матрице Π^* :

$$\Pi^* = \begin{pmatrix} \alpha_1^* & \alpha_1'^* \\ \alpha_2^* & \alpha_2'^* \\ \vdots & \vdots \\ \alpha_{16}^* & \alpha_{16}'^* \\ \beta_1^* & \beta_1'^* \\ \beta_2^* & \beta_2'^* \\ \vdots & \vdots \\ \beta_{16}^* & \beta_{16}'^* \end{pmatrix} \begin{matrix} v_1 v_2 v_6 v_8 \rightarrow v_4 \\ 0000 \rightarrow 0 \\ 0001 \rightarrow 0 \\ \vdots \\ 1111 \rightarrow 0 \\ 1111 \rightarrow 1 \\ 1110 \rightarrow 1 \\ \vdots \\ 0000 \rightarrow 1 \end{matrix} \quad (7)$$

где элементы матрицы Π^* удовлетворяют условию нормировки

$$\alpha_i^* + \alpha_i'^* = 1; \quad \alpha_i^* = \beta_i^*, \quad \alpha_i'^* = \beta_i'^*; \quad i = \overline{1, 16}.$$

Элементы матрицы Π^* можно определить по формулам (8) – (13)

$$\alpha_1^* = \alpha_1'^* = \beta_1^* = \beta_1'^* \quad 1 - \frac{{}^1\pi_{ij}^{(l)2} {}^2\pi_{ij}^{(l)6} {}^6\pi_{ij}^{(l)8} {}^8\pi_{ij}^{(l)10} {}^{\overline{10}}\pi_{ii}^{(l)}}{{}^3\pi_{ii}^{(l)5} {}^5\pi_{ii}^{(l)7} {}^7\pi_{ii}^{(l)9} {}^{\overline{9}}\pi_{ii}^{(l)}} \quad (8)$$

$$\alpha_{16}^* = \alpha_1'^* = \beta_{16}^* = \beta_1'^*; \quad \frac{{}^1\pi_{ij}^{(l)2} {}^2\pi_{ij}^{(l)6} {}^6\pi_{ij}^{(l)8} {}^8\pi_{ij}^{(l)10} {}^{\overline{10}}\pi_{ii}^{(l)}}{{}^3\pi_{ii}^{(l)5} {}^5\pi_{ii}^{(l)7} {}^7\pi_{ii}^{(l)9} {}^{\overline{9}}\pi_{ii}^{(l)}} \quad (9)$$

$$\alpha_2^* = \alpha_2'^* = \beta_2^* = \beta_2'^*; \quad 1 - \frac{{}^1\pi_{ii}^{(l)2} {}^2\pi_{ij}^{(l)6} {}^6\pi_{ij}^{(l)8} {}^8\pi_{ij}^{(l)10} {}^{\overline{10}}\pi_{ii}^{(l)}}{{}^3\pi_{ij}^{(l)5} {}^5\pi_{ii}^{(l)7} {}^7\pi_{ii}^{(l)9} {}^{\overline{9}}\pi_{ij}^{(l)}} \quad (10)$$

$$\alpha_3^* = \alpha_3'^* = \beta_3^* = \beta_3'^*; \quad \frac{{}^1\pi_{ii}^{(l)2} {}^2\pi_{ij}^{(l)6} {}^6\pi_{ij}^{(l)8} {}^8\pi_{ij}^{(l)10} {}^{\overline{10}}\pi_{ii}^{(l)}}{{}^3\pi_{ij}^{(l)5} {}^5\pi_{ii}^{(l)7} {}^7\pi_{ii}^{(l)9} {}^{\overline{9}}\pi_{ij}^{(l)}} \quad (11)$$

$$\alpha_4^* = \alpha_4'^* = \beta_4^* = \beta_4'^*; \quad 1 - \frac{{}^1\pi_{ii}^{(l)2} {}^2\pi_{ii}^{(l)6} {}^6\pi_{ij}^{(l)8} {}^8\pi_{ij}^{(l)10} {}^{\overline{10}}\pi_{ii}^{(l)}}{{}^3\pi_{ii}^{(l)5} {}^5\pi_{ij}^{(l)7} {}^7\pi_{ii}^{(l)9} {}^{\overline{9}}\pi_{ij}^{(l)}} \quad (12)$$

$$\alpha_5^* = \alpha_5'^* = \beta_5^* = \beta_5'^*; \quad \frac{{}^1\pi_{ii}^{(l)2} {}^2\pi_{ii}^{(l)6} {}^6\pi_{ij}^{(l)8} {}^8\pi_{ij}^{(l)10} {}^{\overline{10}}\pi_{ii}^{(l)}}{{}^3\pi_{ii}^{(l)5} {}^5\pi_{ij}^{(l)7} {}^7\pi_{ii}^{(l)9} {}^{\overline{9}}\pi_{ij}^{(l)}} \quad (13)$$

Для удаления точечных помех (1-2 элемента другой яркости) значения вычисленной величины количества информации в элементе изображения сравнивают с порогом H .

Значение порога H определяется из условия:

$$I(v_4 = M_1 | v_1 = v_2 = M_1; v_6 = v_8 = M_2) < H < I(v_4 = M_1 | v_1 = M_1; v_2 = v_6 = v_8 = M_2) \quad (14)$$

Если количество информации в элементе изображения будет больше порога, то значение элемента $v_4^{(l)}$ меняется на противоположное.

Например, для

$${}^1\pi_{ii}^{(l)} = {}^2\pi_{ii}^{(l)} = {}^6\pi_{ii}^{(l)} = {}^8\pi_{ii}^{(l)} = 0,95,$$

если три окрестных элемента имеют знаки противоположные $v_4^{(l)}$, то количество информации в элементе изображения будет составлять

$$I(v_4^{(l)} | v_1^{(l)}, v_2^{(l)}, v_6^{(l)}, v_8^{(l)}) = -\log \left(\frac{{}^1\pi_{ii}^{(l)2} {}^2\pi_{ij}^{(l)6} {}^6\pi_{ij}^{(l)8} {}^8\pi_{ij}^{(l)10} {}^{\overline{10}}\pi_{ii}^{(l)}}{{}^3\pi_{ij}^{(l)5} {}^5\pi_{ii}^{(l)7} {}^7\pi_{ii}^{(l)9} {}^{\overline{9}}\pi_{ij}^{(l)}} \right) = 6,2321 \text{ бит.} \quad (15)$$

Если только два окрестных элемента имеют знаки противоположные $v_4^{(l)}$, то количество информации в элементе изображения равно:

$$I(v_4^{(l)} | v_1^{(l)}, v_2^{(l)}, v_6^{(l)}, v_8^{(l)}) = -\log \left(\frac{{}^1\pi_{ii}^{(l)2} {}^2\pi_{ij}^{(l)6} {}^6\pi_{ij}^{(l)8} {}^8\pi_{ij}^{(l)10} {}^{\overline{10}}\pi_{ii}^{(l)}}{{}^3\pi_{ii}^{(l)5} {}^5\pi_{ij}^{(l)7} {}^7\pi_{ii}^{(l)9} {}^{\overline{9}}\pi_{ij}^{(l)}} \right) = 1,9842 \text{ бит.} \quad (16)$$

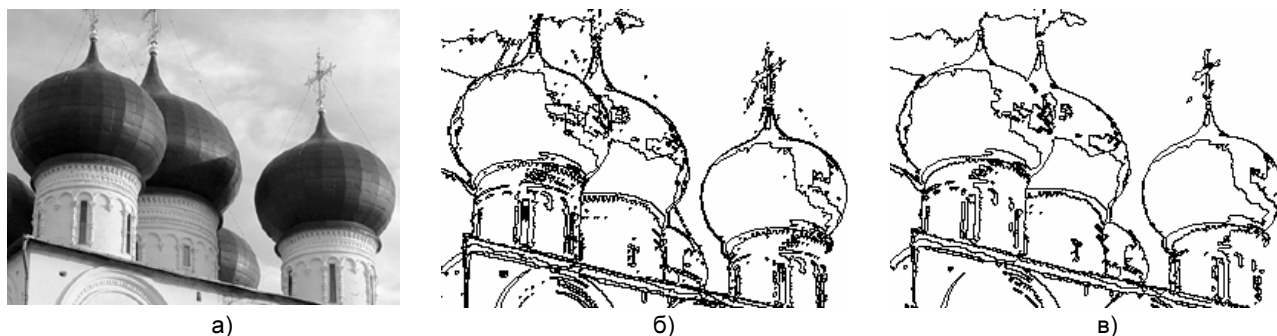


Рис. 6. Исходное ЦПИ (а); изображение с выделенными контурами без удаления (б) и с удалением (в) точечных помех

А значение порога определится из условия (14):
 $1,9842 < H < 6,2123$.

На рис.6 приведены увеличенные фрагменты изображений с выделенными контурами без удаления (б) и с удалением (в) точечных помех.

Как видно из приведенных данных, изображение на рис. 6в имеет более низкий процент ложных контуров, чем изображение на рис. 6б.

Заключение

Разработанный метод выделения контуров в цифровых полутоновых изображениях, основанный на вычислении количества информации в элементах изображения, эффективен с позиции решения поставленной задачи обработки изображений и требует небольших вычислительных затрат.

Литература

1. Новейшие методы обработки изображений. / Под ред. А. А. Потапова. – М.: ФИЗМАТЛИТ, 2008. – 496 с.
2. Гонсалес Р., Вудс Р. Цифровая обработка изображений. – М.: Техносфера, 2005. – 1072 с.
3. Трубин И.С. Метод моделирования цифровых полутоновых изображений / И.С. Трубин, Е.В. Медведева, О.П. Булыгина - «Инфокоммуникационные технологии», Том 6, №1, 2008 – С.94-99.
4. Петров Е.П. Вычисление статистической избыточности статических изображений / Е. П. Петров, Медведева Е.В. - «Вопросы радиоэлектроники», сер. РЛТ, 2008, вып.3 – Москва, 2008. – С.76-83.
5. Фано Р. Передача информации. Статистическая теория связи / Пер. с англ. под ред. Р.Л. Добрушина. – М.: Мир, 1965. – 438 с.

METHOD OF ALLOCATION OF CONTOURS IN THE IMAGE ON THE BASIS OF CALCULATION OF AMOUNT OF INFORMATION

Medvedeva E. V.

The method of allocation of contours in digital half-tone pictures is offered. Digital half-tone pictures are submitted by g -digit binary images. The method is based on calculation of amount of information in the elements of the binary image. For correction of a contour of the image the method of removal of false contours (of dot handicaps) is offered. Results of modeling are presented.

Уважаемые авторы!

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

1) Требования к текстовым материалам и сопроводительным документам:

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

2) Требования к иллюстрациям:

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

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

КАЛИБРОВКА ПОЛОЖЕНИЯ ВИДЕОКАМЕРЫ В ЗАДАЧЕ ОПТИЧЕСКОЙ ЛАЗЕРНОЙ ТРИАНГУЛЯЦИИ

Давыденко Е.В., главный специалист ОАО «Промсвязьбанк», egordv@gmail.com

Приоров А.Л., к.т.н., доцент Ярославского государственного университета им. П.Г. Демидова, andcat@yandex.ru

Ключевые слова: лазерная триангуляция, калибровка видеокамеры, оптическая система, обработка изображений, цифровая модель, сканирование объекта, вычисление положения, алгоритм преобразования.

Введение

Оптическая лазерная триангуляция – один из самых точных и быстрых способов получения цифровых трехмерных моделей реальных объектов [1, 2]. Метод основан на освещении объекта лазерным лучом и регистрации отраженного от объекта излучения с помощью ПЗС-матрицы или иного регистрирующего оборудования (рис. 1).

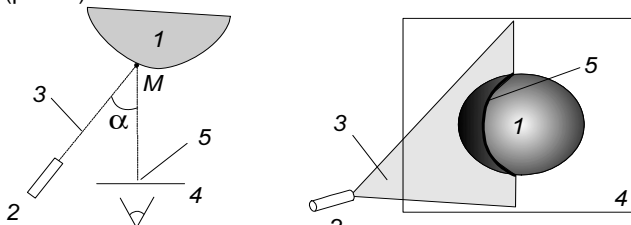


Рис. 1. Принцип работы системы оптической лазерной триангуляции: 1) сканируемый объект; 2) лазер с цилиндрической линзой; 3) лазерный луч; 4) экранная плоскость камеры; 5) проекция лазерного луча на экранную плоскость камеры, повторяющая форму объекта

Принцип действия системы оптической лазерной триангуляции состоит в следующем. Источник лазерного излучения под определенным углом освещает объект лазерным лучом. Изображение объекта с лазерным лу-

Дано описание автоматического метода калибровки видеокамеры для задачи лазерной триангуляции, использующего только изображения, снимаемые видеокамерой. Рассмотрены этапы построения данного метода и анализ его работы в различных условиях. Приведены экспериментально полученные результаты, указаны методы повышения точности калибровки.

чом на его поверхности регистрируется цифровой видеокамерой. Путем анализа данного цифрового изображения вычисляются трехмерные координаты точек его поверхности, на которых присутствует лазерный луч. Сканирование всего объекта возможно путем смещения лазерного луча вдоль всей его поверхности. Примеры моделей, полученных на выходе такой системы, приведены на рис. 2.

Задача калибровки положения видеокамеры

Для корректного вычисления трехмерных координат поверхности сканируемого объекта необходимо знать взаимное расположение лазера и видеокамеры в пространстве. В данной работе предлагается метод вычисления трехмерных координат видеокамеры и лазера без применения специализированных технических средств, с использованием в качестве входных данных только изображений, снимаемых самой видеокамерой. Для позиционирования видеокамеры в сканируемую сцену помещается калибровочный шаблон определенной структуры с нанесенными метками (рис. 3).

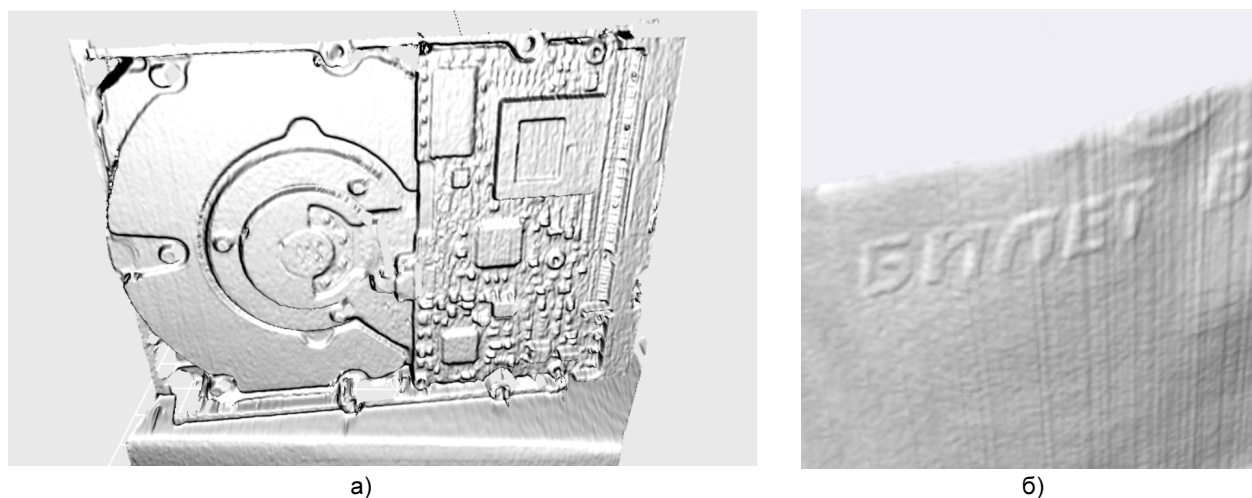


Рис. 2. Пример трехмерных моделей, полученных с применением описываемого алгоритма калибровки: а) модель жесткого диска IBM PC; б) фрагмент трехмерной модели купюры достоинством 500 рублей (тиснение шрифта корректно отсканировано)

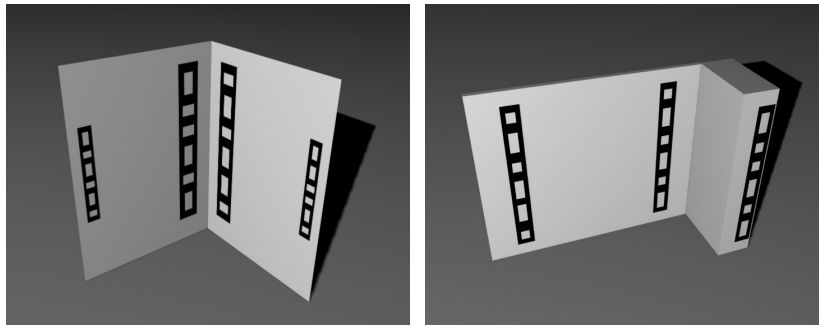


Рис. 3. Модели различных калибровочных шаблонов, используемых в системе, с нанесенными калибровочными полосами

Взаимное расположение меток на калибровочном шаблоне заранее известно, и, автоматически определив положение меток на анализируемом изображении, можно вычислить соответствие между их двумерными координатами на плоскости изображения и трехмерными – на калибровочном шаблоне. Положение лазера также определяется путем автоматического анализа полученного видеокamerой изображения луча лазера на поверхности шаблона.

В качестве калибровочных меток в описываемом алгоритме использованы контрастные калибровочные полосы специальной структуры, показанные на рис. 3. Задача автоматизированного обнаружения калибровочных полос может быть эффективно решена с использованием преобразования Хафа [3] после проведения процедуры детектирования краев в изображении.

На первом этапе работы алгоритма производится детектирование краев в изображении. Анализ экспериментальных данных позволяет сделать вывод о том, что наибольшей эффективностью в рассматриваемой задаче обладают операторы Собеля и Превитта. Детектор краев Кэнни в силу ряда его особенностей показал меньшую эффективность.

После выполнения операции детектирования краев на выходе получается приближенное изображение модуля градиента на изображении. Каждая калибровочная полоса преобразуется в четыре линии, соответствующие ее краям. Данные линии могут быть обнаружены с помощью преобразования Хафа.

Преобразование Хафа представляет собой метод обнаружения особенностей в изображении. В качестве искомой особенности может выступать определенная геометрическая фигура или фрагмент изображения, заданный набором точек. Преобразование заключается в переводе информации из координатного пространства анализируемого изображения в пространство так называемого аккумулятора (аккумуляторного массива). Аккумулятор – некоторое накопительное пространство с размерностью, равной числу параметров для однозначного задания искомой особенности.

Любая прямая на изображении может быть параметризована с помощью двух параметров – удалением от начала координат R и углом наклона θ . В таком случае прямая задается уравнением $R = x \cos(\theta) + y \sin(\theta)$. Следует отметить, что под прямой в данном случае понимается линия неограниченной длины, проходящая через все изображение.

Пусть $I(x, y)$ – входное полутоновое изображение размера M на N точек, и T – некоторый порог. Обозначим аккумуляторный массив преобразования как $A(R, \theta)$. Тогда алгоритм преобразования будет иметь вид:

Для y от 1 до N

Для x от 1 до M

Если $I(x, y) > T$ то

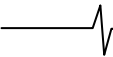
Для θ от 0 до 360

{
 $R = x \cos(\theta) + y \sin(\theta)$;
 $A(R, \theta) = A(R, \theta) + 1$;
 }

Данное преобразование представляет собой последовательный перебор всех точек изображения с проверкой их интенсивности. Если интенсивность точки исходного изображения с координатами (x, y) выше определенного порога, то необходимо увеличить значения тех точек аккумулятора, соответствующие параметрам (R, θ) которых описывают прямые, проходящие через данную точку (x, y) . После преобразования каждая относительно яркая прямая в исходном изображении преобразуется в локальный максимум в координатах аккумулятора, а амплитуда максимума пропорциональна яркости исходной линии.

Далее для поиска прямых в исходном изображении необходимо провести анализ пространства аккумулятора на наличие локальных максимумов достаточной амплитуды. Для поиска локальных максимумов используется следующий итерационный алгоритм.

На первой итерации выбирается начальная точка (x_0, y_0) . В круговой окрестности этой точки радиуса R ищется точка, большая по яркости. Если такая точка найдена, то центр текущей области поиска переносится в эту точку, и алгоритм повторяет свою работу уже в ее окрестности. Если текущая точка является самой яркой в своей окрестности, то алгоритм прекращает свою работу. Первоначально области поиска располагаются через равные промежутки по всей поверхности аккумулятора. После выполнения необходимого числа итераций алгоритм сходится к локальным максимумам в аккумуляторе. Далее составляется список из наиболее ярких максимумов, отсортированный по убыванию амплитуды. В терминах исходного изображения этот список представляет собой список параметров (R, θ) наиболее ярких прямых в исходном изображении.



Далее производится анализ этого списка для нахождения непрерывных прямых с длиной, не меньше заданной. Этот шаг необходим, так как преобразование Хафа дает информацию только о положении и наклоне прямой, и лишь косвенно содержит информацию о ее протяженности, и, кроме того, оно не учитывает разрывы в прямых.

Следующим этапом в анализе результатов преобразования Хафа является отбор пар прямых, которые потенциально могут составлять края одной калибровочной полосы. Для соответствия краям производится проверка каждой возможной пары прямых на выполнение следующих условий:

1. Прямые должны быть расположены близко друг к другу: $|R_2 - R_1| < \Delta R$.

2. Прямые должны быть близки к параллельности: $|\theta_2 - \theta_1| < \Delta \theta$.

3. Длина прямых должна быть примерно одинакова: $|(t_1^E - t_1^S) - (t_2^E - t_2^S)| < \Delta t$.

4. Градиенты прямых должны быть близки к перпендикулярности: $|G_2 - G_1| - 180 < \Delta G$.

5. Параметры начал прямых, так же как и параметры их концов, должны быть примерно равными: $|(t_1^S - t_2^S)| < \Delta t$, $|(t_1^E - t_2^E)| < \Delta t$.

6. Градиенты прямых должны быть направлены не в одну сторону: $|\theta_1 - 90 - G_1| < \Delta \theta^G$.

В данном списке нижние индексы обозначают первую и вторую прямую из пары. Здесь использованы следующие переменные: θ_1, θ_2 – параметры θ (угла наклона) прямых, R_1, R_2 – параметры R (расстояния от начала координат) прямых, G_1, G_2 – средняя величина направления градиента на прямой, t_1^S – параметры начальной точки непрерывного отрезка, t_1^E – параметры конечной точки непрерывного отрезка.

Величины ΔR , Δt , ΔG и $\Delta \theta^G$ – пороговые величины, задающие возможные допуски выполнения требований из приведенного списка. Данные допуски адаптивно снижаются в случае невозможности обнаружить необходимое количество калибровочных полос, и анализ списка производится заново. При невозможности найти прямые после нескольких итераций алгоритм

делает вывод об отсутствии необходимого числа калибровочных полос на изображении или недостаточно высоком его качестве.

После нахождения необходимого количества корректных пар прямых производится считывание кода, нанесенного на центр полосы, составленной найденными прямыми. Код представляет собой чередование черных и белых областей переменной ширины (рис. 4, рис. 5). Для кодирования в двоичном базисе области могут быть единичной и двойной ширины. Единичная ширина принимается за двоичный ноль, двойная ширина – за двоичную единицу.

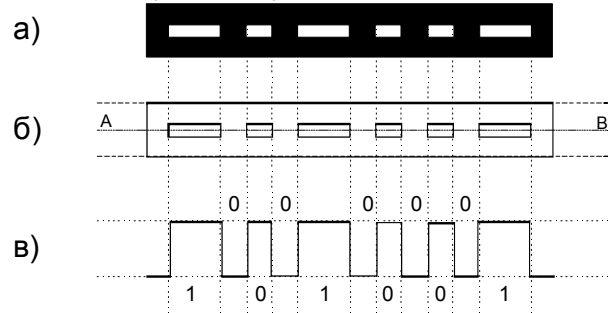


Рис. 4. Процедура считывания кода на калибровочной полосе: а) исходная калибровочная полоса; б) изображение полосы после применения оператора Собеля (изображение инвертировано); в) развертка кода, считанного вдоль прямой АВ

Для считывания кода, пример которого показан на рис. 4а, на первом этапе определяются параметры прямой, лежащей вдоль оси калибровочной полосы (прямая АВ на рис. 4б), и составляется зависимость интенсивности точек изображения от смещения вдоль этой прямой (рис. 4в). Данная зависимость анализируется, и составляется последовательность, представляющая собой зависимость ширины импульса от его порядкового номера. Полученная зависимость передается на алгоритм дискриминации импульсов. В результате составляется двоичная последовательность, в которой импульс единичной ширины принимается за 0, а импульс двойной ширины за 1.

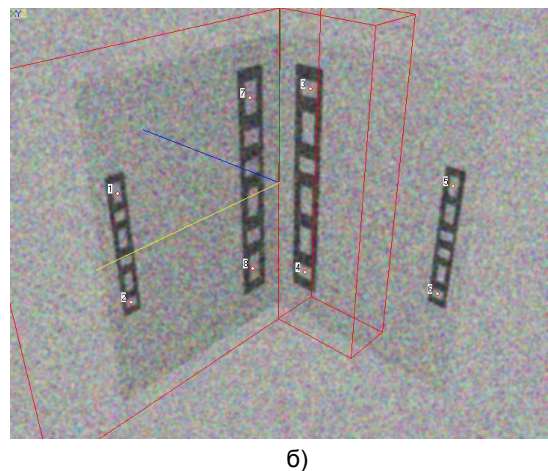
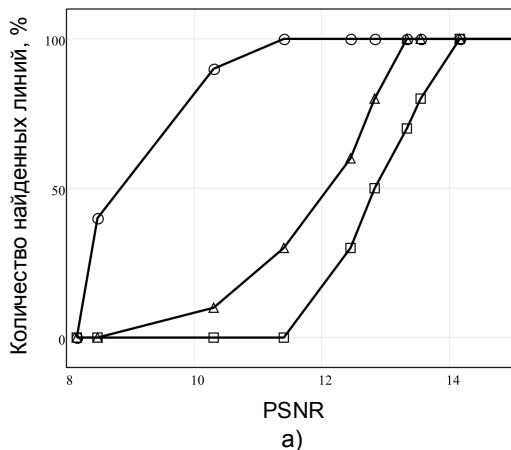


Рис. 5. а) количество найденных линий в процентах от максимально возможного в зависимости от отношения сигнал/шум (PSNR) анализируемого изображения; б) результат калибровки по искаженному изображению (изображение искажалось аддитивным гауссовым шумом с последующим размытием)

Далее эта последовательность проверяется на присутствие в ней заранее заданных возможных кодовых последовательностей. Они выбираются по специализированному базису и различны для различных реализаций калибровочного шаблона. Система анализирует количество найденных калибровочных полос, их расположение и нанесенный штриховой код. Исходя из этой информации, делается вывод о том, какой шаблон в текущий момент используется. Далее информация о положении полос на изображении и тип шаблона передаются в систему вычисления трехмерного положения видеокамеры. Данный метод обнаружения калибровочных полос в силу их протяженности обладает достаточно высокой надежностью и низкой чувствительностью к шумам. Пример калибровки по зашумленным изображениям показан на рис. 5.

Однако одной из основных проблем такого подхода является присутствие перспективных искажений. При наблюдении калибровочной полосы под углом к оптической оси видеокамеры ее форма искажается. Первоначально прямоугольная полоса в таком случае выглядит как трапеция (рис. 6).

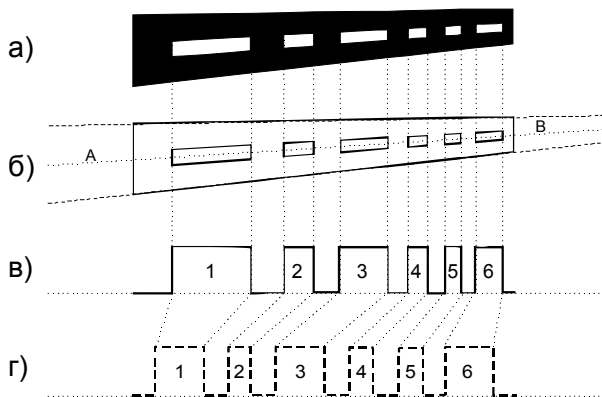
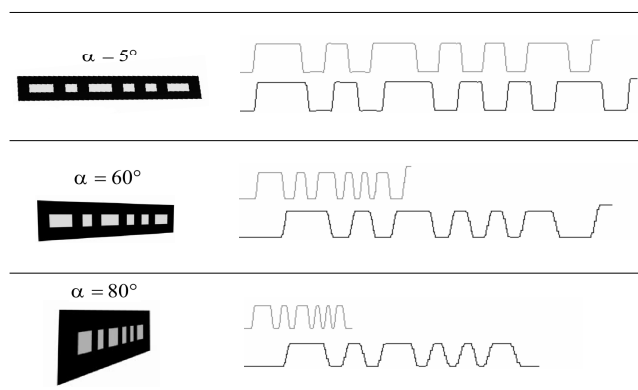
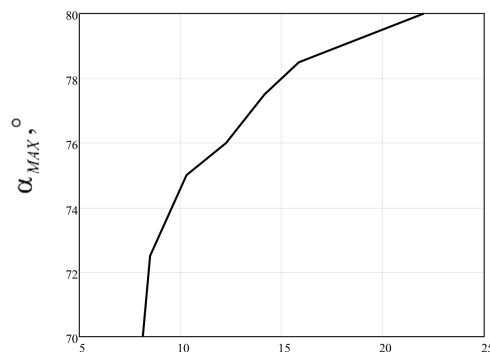


Рис. 6. Пример перспективных искажений кода на калибровочной полосе: а) калибровочная полоса, наблюдаемая с перспективной проекцией; б) результат применения оператора Собеля; в) развертка кода, прямую считанного по прямой АВ; г) код, первоначально содержащийся на неискаженной полосе

Такое искажение не только меняет форму калибровочной полосы, но и в значительной степени искажает код, нанесенный на ней.



а)



б)

Рис. 7. а) результаты работы алгоритма автоматической коррекция интервала дискретизации считывания кодовой последовательности в зависимости от угла наблюдения полосы α . Развертка кода на линии без коррекции обозначена серым цветом (вверху), после коррекции – черным (внизу); б) зависимость максимально возможного угла наблюдения полосы α_{MAX} от отношения сигнал/шум анализируемого изображения

В рассматриваемой системе данный тип искажений в необходимой степени минимизируется путем использования специального алгоритма автоматического изменения интервала дискретизации считывания кода (рис. 7).

На первом этапе данный алгоритм анализирует степень отклонения формы калибровочной полосы от идеального прямоугольника и по этим данным определяет величину перспективных искажений. В случае, когда перспективные искажения отсутствуют, интервал дискретизации на всем протяжении калибровочной полосы равен 1, и чтение кода на калибровочной полосе производится с шагом в 1 пиксель. В случае, когда линии не параллельны, единичный интервал дискретизации выбирается на широком конце полосы и уменьшается при приближении к узкому концу. Интервал дискретизации меньше 1 пикселя соответствует переходу на субпиксельное разрешение. Для реализации такой выборки точек на изображении с нецелой координатой алгоритм использует взвешенное окно 2x2 пикселя.

В результате применения данного алгоритма искажения, обусловленные эффектом перспективы, минимизируются и практически не влияют на надежность работы алгоритма считывания кода.

Вычисление положения видеокамеры

После определения положения меток на исходном изображении необходимо вычислить вектор $\mathbf{X} = (x_0, y_0, z_0, \omega, \varphi, \kappa, c, s)$ параметров видеокамеры в соответствии с моделью перспективной проекции. Данный вектор включает в себя не только координаты видеокамеры x_0, y_0, z_0 и углы Эйлера ω, φ, κ ее ориентации, но и фокусное расстояние видеокамеры c и коэффициент диспропорции изображения s . При необходимости в данный вектор также могут быть добавлены параметры, описывающие радиальную и другие виды дисторсии изображения. Таким образом, используя только одно изображение на входе блока калибровки, можно определить не только требуемые внешние, но и внутренние параметры видеокамеры. Определение данного набора параметров с помощью более простых методов калибровки, например путем однократной калибровки по плоскому регулярному шаблону, в общем случае невозможно.

Для вычисления элементов вектора \mathbf{X} необходимо определить соответствие между двумерными координатами (ξ, η) положения меток на экранной плоскости видеокамеры и их трехмерными координатами (x, y, z) на шаблоне [4, 5]. В качестве калибровочных меток выбираются центры областей, составляющих код на калибровочной полосе.

Для нахождения элементов вектора \mathbf{X} необходимо минимизировать модули расстояний между координатами меток на изображении и их расчетными значениями. Данная минимизация эффективно выполняется с помощью итерационного алгоритма Левенберга-Марквардта [6]. Он представляет собой алгоритм нелинейной оптимизации и эффективно работает в случае больших невязок, обладая быстрой сходимостью. Для решения задачи минимизации вводится функция невязок вида:

$$f(\mathbf{X}) = \frac{1}{2} \sum_{j=1}^m \mathbf{r}_j^2(\mathbf{X}). \quad (1)$$

Вектор невязок \mathbf{r} отклонений полученных положений маркеров $\xi(x_i)$ и $\eta(x_i)$ от их расчетных значений $\xi'(x_i)$ и $\eta'(x_i)$ в (1) равен $r_j = |\xi(x_i) - \xi'(x_i)|$, $r_{j+1} = |\eta(x_i) - \eta'(x_i)|$. Задача нахождения минимума функции $f(\mathbf{X})$ методом Левенберга-Марквардта может быть решена итерационно, используя следующее выражение для нахождения величины шага на каждой итерации:

$$\mathbf{X}_{i+1} = \mathbf{X}_i - (\mathbf{H} + \lambda \mathbf{I})^{-1} \nabla f(\mathbf{X}_i). \quad (2)$$

В этом выражении \mathbf{H} – матрица Гессе, вычисленная в точке \mathbf{X}_i . Выражение (2) является ключевым в методе Левенберга-Марквардта и представляет собой комбинацию метода градиентного спуска и метода Гаусса-Ньютона. При малом значении λ алгоритм стремится к методу Гаусса-Ньютона, при большом – к методу градиентного спуска, позволяя, таким образом, одновременно использовать достоинства обоих методов.

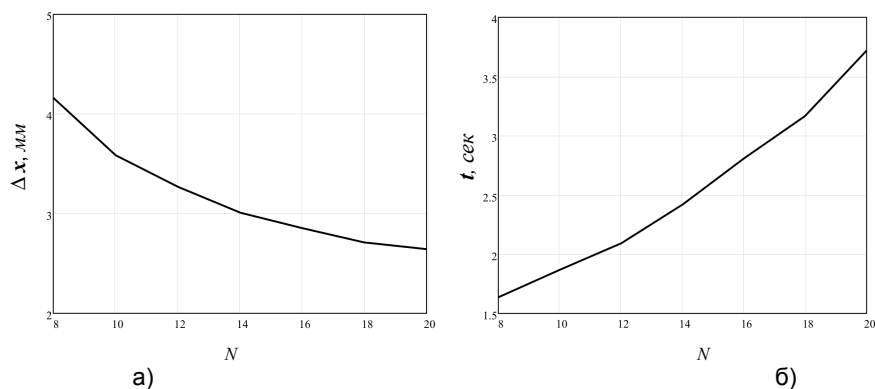


Рис. 8. а) ошибка определения положения видеокамеры Δx в зависимости от количества калибровочных точек N ; б) время калибровки t в зависимости от N

Литература

1. Твердохлеб П.Е., Коронкевич В.П., Косцов Э.Г. и др. 3D лазерные информационные технологии. – Новосибирск: Рос. акад. наук, Сиб. отд-ние, Ин-т автоматки и электротри, 2003.
2. Дунин-Барковский И.И. Построение системы контроля размеров крупногабаритных деталей на основе 3D-системы технического зрения // Измерительная техника. – 2004. № 12. С. 19.
3. Компьютерное зрение / Л. Шапиро, Дж. Стокман; Пер. с англ. – М.: БИНОМ. Лаборатория знаний, 2006.
4. Demirdjian D., Zisserman A., Horaud R. Stereo autocalibration from

Заключение

Погрешность предлагаемого метода калибровки положения видеокамеры составляет в среднем менее 1% от поперечника калибровочного шаблона (при условии, что шаблон занимает около 80% площади изображения) и в первую очередь зависит от точности изготовления шаблона.

При неточном расположении калибровочных полос ошибку можно снизить, увеличив количество точек, по которым производится калибровка. При увеличении количества точек растет и время, необходимое на вычисление положения видеокамеры. Данные зависимости показаны на рис. 8. Система лазерной триангуляции, построенная с применением описанной процедуры калибровки, требует минимального набора компонентов (лазера, видеокамеры и ЭВМ общего назначения) и позволяет отказаться от дополнительного дорогостоящего оборудования (координатно-измерительных машин и т.п.). Форма калибровочного шаблона может быть изменена для учета различных пропорций сканируемых объектов без необходимости перестройки системы. Применение меток для определения положения видеокамеры в свою очередь облегчает процедуру построения полного трехмерного образа объекта методом совмещения информации, полученной сканированием с нескольких ракурсов, а так же позволяет упростить процедуры ее последующей обработки.

Предлагаемый метод калибровки положения видеокамеры применим не только для задачи лазерной триангуляции, но и для других задач, в которых необходимо определение положения и ориентации видеокамеры с использованием только изображений, снимаемых ею. Метод эффективно работает в условиях низкого качества изображений (которые можно улучшить методами, описанными в [7]), значительных перспективных искажений и присутствия естественного фона позади калибровочного шаблона.

one plane // In proc. of 6th european conference on computer vision (ECCV), Dublin, Ireland, 2000. Part 2, P. 625-639.

5. Zhang Z. A flexible new technique for camera calibration // IEEE Trans. on pattern analysis and machine intelligence. 2000. V. 22, Is. 11. P. 1330-1334.
6. Vetterling W.T. Numerical Recipes in C. – New York: Cambridge University Press (Massachusetts), 1992.
7. Цифровая обработка телевизионных и компьютерных изображений / А.В. Дворкович, В.П. Дворкович, Ю.Б. Зубарев и др. – М.: МЦНТИ, 1997.

УДК 004.932

НОРМАЛИЗАЦИЯ КОСМИЧЕСКИХ ИЗОБРАЖЕНИЙ ЗЕМЛИ НА ОСНОВЕ ИХ СОПОСТАВЛЕНИЯ С ЭЛЕКТРОННЫМИ КАРТАМИ

Егошкин Н.А., к.т.н., с.н.с. НИИ «Фотон» Рязанского государственного радиотехнического университета (РГРТУ), gislab@org.etr.ru

Еремеев В.В., д.т.н., директор НИИ «Фотон» РГРТУ, foton@rgta.ryazan.ru

Козлов Е.П., м.н.с. НИИ «Фотон» РГРТУ

Ключевые слова: дистанционное зондирование Земли, геостационарный спутник, нормализация изображений, электронная карта, автоматическое совмещение, координатная привязка

Введение

В практике дистанционного зондирования Земли (ДЗЗ) в последние годы стали внедряться геостационарные космические системы: MeteoSat (Европа), GOES (США), FY (Китай) и др. В настоящее время готовится к запуску российский геостационарный спутник «Электро-Л», который планируется включить в международную космическую систему наблюдения всей земной поверхности. Геостационарные спутники выводятся на орбиту, находящуюся в плоскости экватора, на высоту порядка 35000 км. При этих условиях спутник «зависает» на заданной долготе и обеспечивает периодическую съемку одной и той же поверхности всего земного диска на фоне окружающего космоса (рис. 1). Погрешности измерений пространственного положения и углов ориентации спутника приводят к тому, что изображение диска Земли искажается, прежде всего смещается в регистрационной системе координат, поворачивается и изменяет масштаб. Стоит задача с использованием тех или иных технологий выполнить нормализацию изображения, т.е. привести его к стандартным положению и размерам.

Известные технологии нормализации изображений по звездному небу и наземным ориентирам [1] не обеспечивают необходимой точности, оперативности и надежности решения рассматриваемой задачи. В последние годы появились векторные электронные карты различного масштаба, и делаются активные попытки их использования для координатной обработки данных ДЗЗ, полученных от полярно-орбитальных космических аппаратов, которые осуществляют съемку Земли, вращаясь вокруг нее. Но полностью автоматических технологий решения этой задачи еще не создано [2-4]. Основная трудность состоит в постоянном изменении географического района съемки, условий наблюдения, в выборе устойчивых во времени наземных ориентиров.

Задача нормализации изображений от геостационарных спутников значительно упрощается: они осуществляют съемку одного и того же района земной поверхности; на снимках четко отображаются береговые линии морей, крупных озер и островов, которые устойчивы во времени и представлены в соответствующих электрон-

Рассматриваются алгоритмы и технология нормализации изображений от геостационарных спутников на основе двух подходов. Первый базируется на автоматическом совмещении уникальных фрагментов береговых линий снимка с бинарными масками, формируемыми по электронным картам. Вторым подходом основан на эталонном описании контура диска Земли в специально введенной плоскости и сопоставлении этого описания с реально наблюдаемым на снимке. Предложены алгоритмы повышения точности и надежности решения этой задачи.

ных картах. Кроме этого, снимки от геостационарных спутников (по отношению к полярно-орбитальным) предоставляют уникальную возможность их координатной привязки по контурным точкам диска Земли, которые контрастно и неизменно отображаются на фоне окружающего космоса.

Все это создает предпосылки для создания полностью автоматической технологии высокоточной нормализации космических изображений от геостационарных спутников по электронным картам и контурным точкам диска Земли.

Общая технология координатной привязки

В настоящей работе рассматривается технология надежной координатной привязки, предполагающая реализацию трех основных процессов. Во-первых, на снимке выделяются опорные объекты, которые имеют уникальную геометрическую форму и устойчивы во времени. Для этих объектов на основе ЭК формируются бинарные маски, соответствующие по яркости и геометрической форме аналогичным объектам снимка. Во-вторых, выполняется корреляционное совмещение бинарных масок с одноименными объектами снимка, на основе чего определяются функции геометрического соответствия снимка и карты. В-третьих, на изображении контурные точки диска Земли преобразуются во вспомогательную плоскость. В этой плоскости они образуют эталонный эллипс, который рассматривается как дополнительный элемент ЭК. На основе сопоставления реально наблюдаемого контура с его эталонным описанием решается задача предварительного определения функций геометрического соответствия снимка и карты, что позволяет значительно сузить зоны поиска одноименных фрагментов снимка и карты. Рассмотрим конкретное содержание этих этапов.

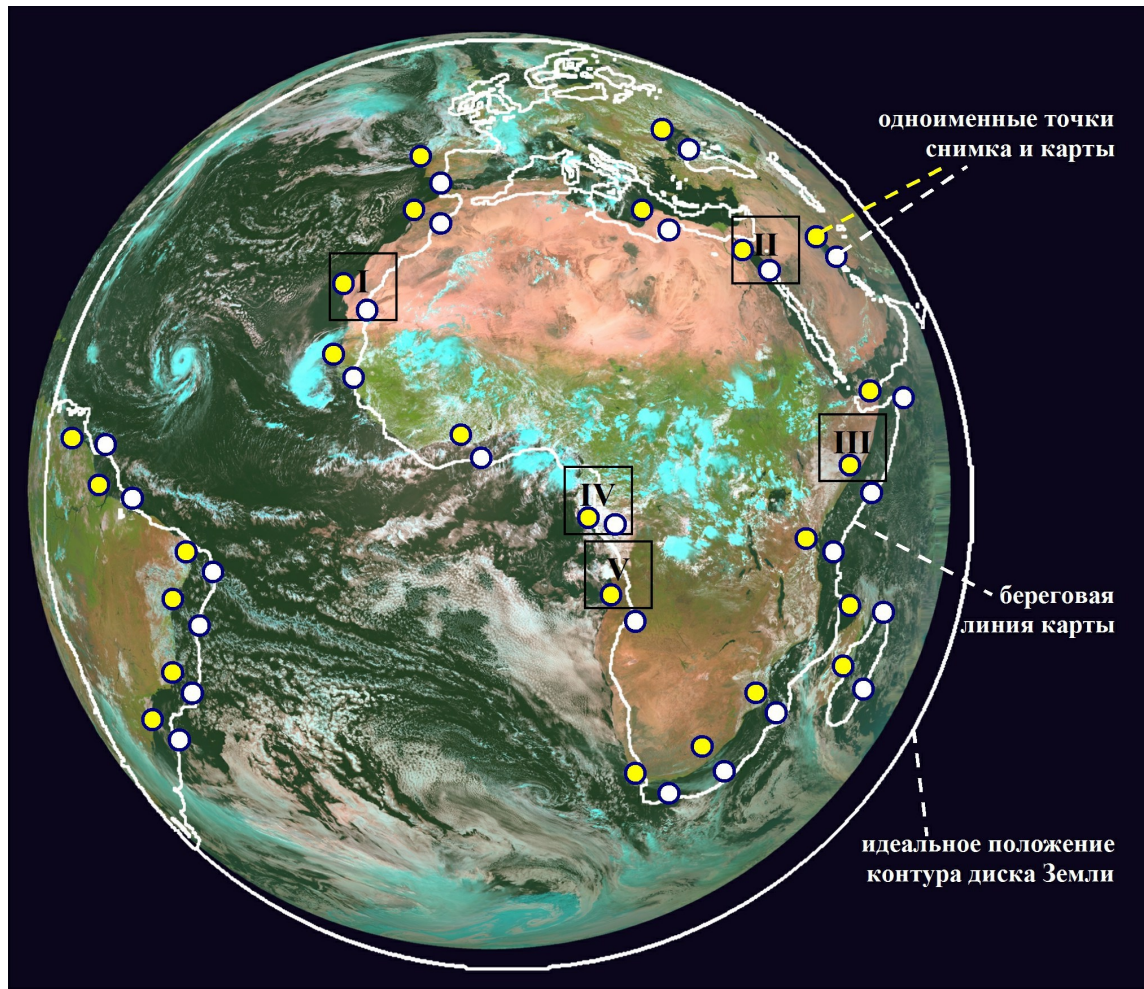


Рис. 1. Изображение диска Земли на фоне окружающего космоса и электронной карты

Формирование бинарных масок. На наблюдаемую часть земного диска из электронной карты отбираются уникальные опорные объекты, соответствующие объектам снимка (рис. 2, слева; фрагмент I на рис. 1), в виде изломов береговых линий и преобразуются из векторного формата в растровый. Степень уникальности опорных фрагментов $E_i = \{(x_j, y_j); j = \overline{1, J}\}, i = \overline{1, I}$, определяется среднеквадратичным отклонением $\sigma_i, i = \overline{1, I}$, береговой линии, содержащей точки $(x_j, y_j), j = \overline{1, J}$, от прямой $y = k_i x + b_i$:

$$\sigma_i = \sqrt{\frac{1}{J} \sum_{j=1}^J ((y_j - (b_i + k_i x_j))^2 + (x_j - (b_i + k_i y_j))^2)}, \quad (1)$$

где $k_i, b_i, i = \overline{1, I}$, находятся из условий

$$\sum_j (b_i + k_i y_j - x_j)^2 = \min \text{ и } \sum_j (b_i + k_i x_j - y_j)^2 = \min.$$

Эталонное положение фрагментов определяется их центрами $(x_{0i}, y_{0i}), i = \overline{1, I}$. Для каждого $E_i, i = \overline{1, I}$, формируется бинарная маска $A_i, i = \overline{1, I}$, в виде двух областей, линия соприкосновения которых совпадает с электронной картой; яркости (нулевая и единичная) соответствуют расположению водной и земной поверхностей на снимке (рис. 2, справа). Для каждого опорного фрагмента формируется набор бинарных масок разной ширины для увеличения надежности совмещения фрагмента снимка и маски. Затем с помощью фрагментов

$E_i, i = \overline{1, I}$, с центрами $(x_{0i}, y_{0i}), i = \overline{1, I}$, формируются зоны поиска в виде расширенных фрагментов снимка $B_i, i = \overline{1, I}$, с очертаниями береговых линий E_{Bi} (см. рис. 2). На рис. 2 и далее индекс i для простоты опущен.

Корреляционное совмещение снимка с картой. С учетом бинарного представления маски A при нахождении коэффициента корреляции

$$\rho = (\overline{BA} - \overline{B} \overline{A}) / \sqrt{D_B D_A} \quad (2)$$

не требуется вычислять среднее значение произведений одноименных элементов изображения и маски \overline{BA} , средних яркостей $\overline{A}, \overline{B}$ и дисперсий D_A, D_B , и (2) можно представить как

$$\rho = \frac{\overline{B_1} - \overline{B_0}}{\sqrt{D_B}} \cdot \frac{\sqrt{q_1 q_0}}{q_1 + q_0}, \quad (3)$$

где $\overline{B_1}, \overline{B_0}$ и q_1, q_0 - средние значения яркостей снимка и число точек под единичной и нулевой областями маски. При $q_1 = q_0$ вычисление ρ еще более упрощается:

$$\rho = 0.5(\overline{B_1} - \overline{B_0}) / \sqrt{D_B}. \quad (4)$$

В результате совмещения карты и снимка формируется множество одноименных точек $O = \{(x_i, y_i), (x_i^*, y_i^*); i = \overline{1, I}\}$, на основе которого по МНК определяются полиномы P_x, P_y , устанавливающие геометрическое соответствие снимка и карты.

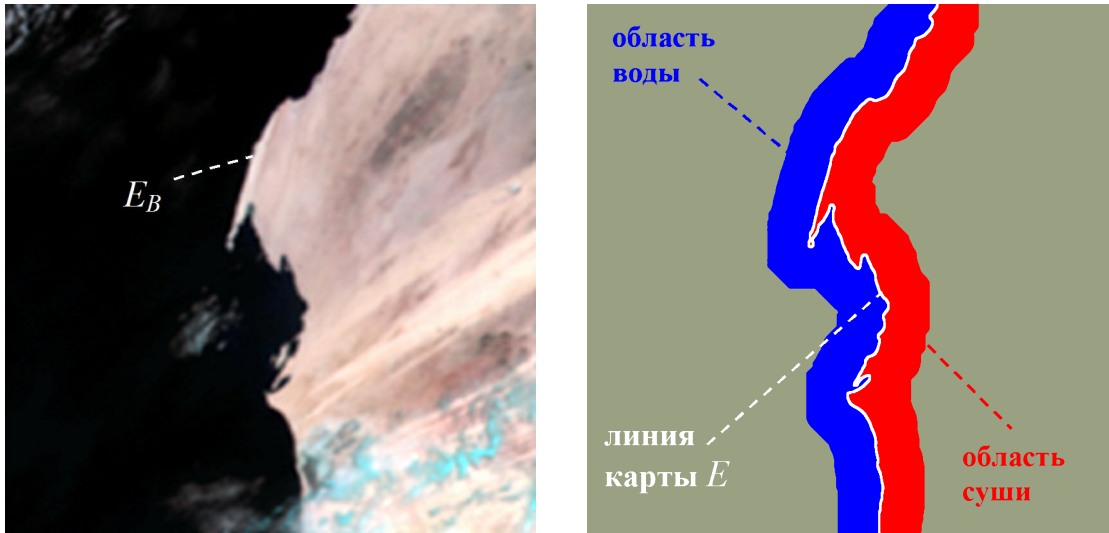


Рис. 2. Уникальный фрагмент снимка (слева) и бинарная маска (справа)

Коэффициенты этих полиномов находятся из условий:

$$\sum_i [P_x(x_i^*, y_i^*) - x_i]^2 = \min, \quad \sum_i [P_y(x_i^*, y_i^*) - y_i]^2 = \min. \quad (5)$$

Коэффициенты полиномов определяют смещение диска Земли по отношению к карте, масштабы по осям, поворот и другие искажения.

Координатная привязка по диску Земли. Выполняется сопоставление эталонного описания контура диска Земли с реально наблюдаемым. В идеальном случае в плоскости, которая перпендикулярна к линии, соединяющей центр Земли и центр масс спутника, крайние точки диска Земли (\tilde{X}, \tilde{Y}) образуют эллипс, а в реальных условиях занимают другое положение (X, Y) . Можно показать, что эталонный эллипс при номинальных параметрах съемки определяется как $\tilde{X}^2/a^2 + \tilde{Y}^2/b^2 = 1$,

$$a = a_3 \sqrt{\frac{X_0^2 b_3^2 + Y_0^2 b_3^2 + Z_0^2 a_3^2}{X_0^2 b_3^2 + Y_0^2 b_3^2 + Z_0^2 a_3^2 - a_3^2 b_3^2}},$$

$$b = \sqrt{\frac{X_0^2 b_3^4 + Y_0^2 b_3^4 + Z_0^2 a_3^4}{X_0^2 b_3^2 + Y_0^2 b_3^2 + Z_0^2 a_3^2 - a_3^2 b_3^2}};$$

(X_0, Y_0, Z_0) – координаты спутника; a_3, b_3 – экваториальный и полярный радиусы земного эллипсоида. Для реального эллипса оцениваются координаты центра $(X_{ц}, Y_{ц})$, размеры полуосей a_d, b_d и угол их наклона φ . Затем устанавливается геометрическое соответствие одноименных точек:

$$\tilde{X} = [(X - X_{ц}) \cos \varphi + (Y - Y_{ц}) \sin \varphi] a / a_d,$$

$$\tilde{Y} = [(Y - Y_{ц}) \cos \varphi - (X - X_{ц}) \sin \varphi] b / b_d. \quad (7)$$

Формулы (7), уточняя положение и размеры диска Земли, позволяют существенно сузить зоны поиска опорных объектов карты на снимке.

Повышение точности и надежности оценки параметров нормализации

При решении рассматриваемой задачи мешающее

действие оказывают такие факторы, как: наличие облучности, частично или полностью закрывающей опорные фрагменты береговых линий; неполное соответствие геометрии береговой линии на снимке и электронной карте; неравномерное распределение уникальных фрагментов по полю изображения. Рассмотрим ряд конструктивных решений, направленных на нейтрализацию действия этих факторов.

Построение ассиметричных масок. Исследование поведения ρ в окрестности области совпадения фрагмента снимка с маской показало, что при различной степени их геометрического сходства

$$\rho = \frac{1 - (\alpha + \beta)}{\sqrt{1 + 2(D_0 + D_1)/(\bar{B}_1 - \bar{B}_0) + 2(\alpha - \beta)(D_1 - D_0)/(\bar{B}_1 - \bar{B}_0)^2 - (\alpha - \beta)^2}}, \quad (8)$$

где α и β – доля элементов суши под нулевой областью маски и доля элементов водной поверхности под единичной областью («чужие» элементы), D_0 и D_1 – выборочные дисперсии изображения под нулевой и единичной областями бинарной маски.

Из (2) и (8) следуют два важных вывода. Во-первых, ρ зависит как от общего уровня «чужих» элементов, доля которых равна $(\alpha + \beta)$, так и от соотношения между α и β , определяемого в знаменателе членами с разностью $(\alpha - \beta)$. Во-вторых, член $\bar{B}A$ в (2), который должен определять меру сходства фрагмента снимка и маски, зависит только от β . $\bar{B}A$ достигает максимального значения $0,5 B_1$ при $B_1 > B_0$ и $\beta = 0$, когда элементы водной поверхности полностью отсутствуют под единичной областью маски. В результате центрирования и нормирования $\bar{B}A$ формируется коэффициент корреляции, числитель которого уже зависит от суммы $(\alpha + \beta)$, а знаменатель от разности $(\alpha - \beta)$; в результате точка наилучшего сходства в общем случае смещается.

Подобного рода ошибки могут быть устранены путем использования ассиметричных бинарных масок. В этом случае выражение (8) принимает следующий вид:

$$\rho = \frac{(1 - \alpha - \beta) \sqrt{q_0 q_1 / (q_0 + q_1)^2}}{\sqrt{\pi(1 - \pi) + \pi D_0 / (\bar{B}_1 - \bar{B}_0)^2 + (1 - \pi) D_1 / (\bar{B}_1 - \bar{B}_0)^2}}, \quad (9)$$

где $\pi = \beta + q_0 (1 - \alpha - \beta) / (q_0 + q_1)$.

Максимальное значение ρ достигается при

$$q_0 = \frac{q_0 + q_1}{1 + \sqrt{\frac{D_0 + \alpha(\overline{B}_1 - \overline{B}_0)^2}{D_1 + \beta(\overline{B}_1 - \overline{B}_0)^2}}}. \quad (10)$$

Только при полном совпадении E с E_B , т.е. при $\alpha = \beta = 0$, ошибка отсутствует, и $\rho = \rho_{\max}$ при несимметричной маске с $q_0 = (q_0 + q_1) \sqrt{D_1} / (\sqrt{D_0} + \sqrt{D_1})$, либо для симметричной маски ($q_0 = q_1$) при равенстве дисперсий $D_0 = D_1$.

Компенсация систематических ошибок. Такие ошибки обусловлены несовпадением береговых линий на снимке и карте (рис. 3, фрагмент II на рис. 1).



Рис. 3. Несовпадение береговой линии на снимке и карте

Ошибки определения ОТМ, связанные с несовпадением береговой линии на снимке и карте, от сеанса к сеансу съемки повторяются и устраняются с помощью следующего алгоритма.

По описанной выше технологии для изображения $B(x, y)$ формируются полиномы геометрического соответствия снимка и карты P_x, P_y . Далее в полиномы P_x, P_y вводится серия искажений $\Omega_{xs}, \Omega_{ys}, s = \overline{1, S}$, определяющих различные повороты и смещения изображения по отношению к ЭК. В результате формируется набор полиномов $P_{xs} = P_x + \Omega_{xs}, P_{ys} = P_y + \Omega_{ys}, s = \overline{1, S}$. По этим полиномам изображение $B(x, y)$ преобразуется в серию изображений $B_s(x, y), s = \overline{1, S}$; при этом P_x, P_y «работают» на совмещение одноименных фрагментов снимка и карты, а $\Omega_{xs}, \Omega_{ys}, s = \overline{1, S}$, вносят рассогласования. Затем $B_s(x, y), s = \overline{1, S}$, совмещаются с ЭК, в результате чего определяются центральные точки фрагментов изображения $(x_{is}, y_{is}), i = \overline{1, I}, s = \overline{1, S}$. Эти точки преобразуются по обратным полиномам $\Omega_{xs}^{-1}, \Omega_{ys}^{-1}, s = \overline{1, S}$, в плоскость изображения $B(x, y)$, образуя набор точек $(x_{is}^*, y_{is}^*), i = \overline{1, I}, s = \overline{1, S}$. В результате формируются смещения $\Delta x_{is} = x_{is} - x_{is}^*, \Delta y_{is} = y_{is} - y_{is}^*, i = \overline{1, I}, s = \overline{1, S}$, их средние значения

$$\Delta x_i = \frac{1}{S} \sum_{s=1}^S \Delta x_{is}, \Delta y_i = \frac{1}{S} \sum_{s=1}^S \Delta y_{is}$$

и СКО

$$\sigma_{xi} = \sqrt{\frac{1}{S} \sum_{s=1}^S \Delta x_{is}^2 - (\Delta x_i)^2},$$

$$\sigma_{yi} = \sqrt{\frac{1}{S} \sum_{s=1}^S \Delta y_{is}^2 - (\Delta y_i)^2}, i = \overline{1, I}.$$

Величины $\Delta x_i, \Delta y_i, i = \overline{1, I}$, характеризуют систематические ошибки по осям x и y , обусловленные несовпадением по геометрии береговых линий снимка и карты, и могут служить поправками к координатам $x_i, y_i, i = \overline{1, I}$. Величины σ_{xi} и $\sigma_{yi}, i = \overline{1, I}$, характеризуют случайные составляющие ошибки совмещения одноименных фрагментов. Экспериментально установлено, что СКО оценки систематических ошибок не превышает 0.2 пикселя.

Учет уникальности опорных фрагментов. Уникальные фрагменты карты неравномерно расположены по снимку, а их число ограничено и измеряется несколькими десятками. Ставится задача расширить набор опорных фрагментов за счет тех, которые содержат почти прямолинейные береговые линии (рис. 4, слева; фрагмент III на рис. 1).

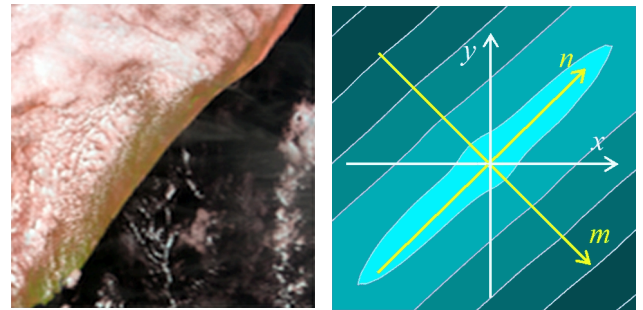


Рис. 4. Фрагмент снимка с линейной береговой линией (слева) и его АКФ

Установлено, что изолинии автокорреляционной функции (АКФ) опорных фрагментов имеют форму, близкую к эллипсу, направления осей m и n которого определяются ориентацией береговой линии (рис. 4, справа). Этот факт дает основание учитывать направления наилучшего (ось m) и наихудшего (ось n) совмещения для всех опорных фрагментов и к найденным ОТМ относиться с разной степенью доверия по осям m и n . В этом случае геометрическое соответствие снимка и карты находится из условий:

$$\sum_{i=1}^I (K_{ni} \Delta n_i^2) = \min, \sum_{i=1}^I (K_{mi} \Delta m_i^2) = \min, \quad (11)$$

где K_{ni} и K_{mi} - весовые коэффициенты, определяющие различную степень доверия по направлениям m и n для i -го фрагмента,

$$\Delta n_i = [P_x(x_i, y_i) - x_i] \cos \alpha_i - [P_y(x_i, y_i) - y_i] \sin \alpha_i,$$

$$\Delta m_i = [P_x(x_i, y_i) - x_i] \sin \alpha_i + [P_y(x_i, y_i) - y_i] \cos \alpha_i, \alpha_i - \text{угол между системами координат } xy \text{ и } mn, i = \overline{1, I}.$$

Углы ориентации АКФ и параметры «округлости» можно найти методом главных компонент. Для этого определяются собственные числа $\lambda_j, j = \overline{1, 2}$, $\lambda_1 \geq \lambda_2$, и собственные вектора e_{ij} матрицы вторых моментов. После этого осуществляется переход к ортогональным некоррелированным координатам m и n :

$$\begin{bmatrix} n \\ m \end{bmatrix} = \begin{bmatrix} e_{11} & e_{12} \\ e_{21} & e_{22} \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix}. \quad (12)$$

Отношение λ_1/λ_2 характеризует, насколько изолиния АКФ в форме эллипса приближается к окружности. Конкретные значения коэффициентов остроконечности АКФ вдоль осей m и n , используемые в (11), находятся как

$$K_m = (\varepsilon_m + 3)\sigma_m \lambda_2, \quad K_n = (\varepsilon_n + 3)\sigma_n \lambda_1, \quad (13)$$

где ε_m , ε_n и σ_m , σ_n – эксцессы и СКО, найденные в окрестности максимума АКФ и позволяющие сравнить остроту вершин срезов АКФ вдоль осей m и n .

В конечном счете алгоритм дает возможность использовать для координатной привязки фрагменты, содержащие линейные участки береговых линий. Экспериментально установлено, что за счет этого набор опорных фрагментов расширяется примерно на 30%, что позволяет более плотно покрыть всю область определения снимка и снизить погрешность его совмещения с картой до 0.2 - 0.4 пикселя.

Идентификация облачности на водных участках. Наличие облачности в зоне поиска (рис. 5, сверху; фрагмент IV на рис. 1) приводит к ошибкам совмещения снимка и карты.

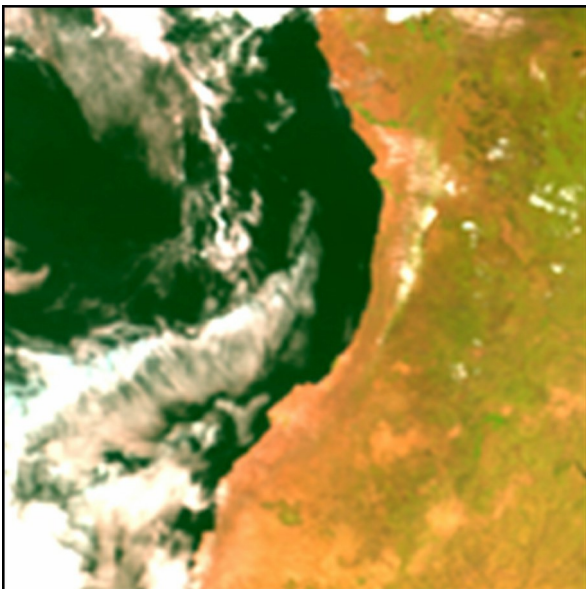


Рис. 5. Устранение облачности на водных участках

Нейтрализация этого фактора основана на двух предпосылках. Во-первых, зоны поиска содержат единственную береговую линию, разделяющую воду и сушу, которая от сеанса к сеансу съемки может смещаться в двух направлениях. Во-вторых, облачные образования представляют собой связанные области, яркость которых всегда выше яркости пикселей водной поверхности.

С учетом этого в зоне поиска определяются три области (рис. 5, снизу): L – область гарантированного участка суши; W – область гарантированного участка водной поверхности; U – область, непосредственно прилегающая к береговой линии и содержащая как участки суши, так и участки воды. Области строятся на основе знаний о допустимых диапазонах смещения береговой линии по осям x и y . На основе анализа гистограммы в области U решается задача сегментации на две подобласти: U_0 – элементы водной поверхности и U_1 – элементы суши. Далее формируется объединенная область элементов водной поверхности $W^* = W \cup U_0$ и находятся элементы, принадлежащие облачным образованиям, окруженные элементами W^* .

В результате этих операций области водной поверхности, закрытые облаками, восстанавливаются (рис. 5, снизу) при сохранении береговых линий. Экспериментально установлено, что алгоритм позволяет на 20-30% снизить число ложных совмещений, вызванных частичным закрытием облаками водных участков опорных фрагментов.

Отбраковка участков береговых линий, закрытых облаками. Закрытый облаками участок береговой линии, составляющий часть фрагмента, отбраковывается при $q_{обл,0}/q_0 > K_{обл}$, где $q_{обл,0}$ – число облачных точек, q_0 – число нулевых точек бинарной маски, $K_{обл}$ – пороговое значение доли облачности. Определение $q_{обл,0}$ производится на основе пороговой сегментации изображения $G(x, y)$, находящегося под бинарной маской. Порогом является медианное значение яркости $med G(x, y)$, которое отделяет водную поверхность под нулевой частью маски от облачных образований.

Отбраковка участков с несовпадающими береговыми линиями снимка и карты. Выполнение этой операции основано на многокритериальном анализе наличия облачных образований и несовпадений береговых линий снимка и карты на отдельных участках опорного фрагмента и принятия решения об их использовании или отбраковке. В качестве критериев используются статистические характеристики водной и земной поверхностей (средняя яркость, дисперсия, АКФ и др.) и степень закрытия участка облаками. Бинарная маска разбивается на L участков $A_i = \{A_i^1, A_i^2, \dots, A_i^l\}$, $i = \overline{1, I}$ (рис. 6; фрагмент V на рис. 1).

В процессе совмещения A_i^l , $i = \overline{1, I}$, $l = \overline{1, L}$, с соответствующим участком снимка B_i^l , $i = \overline{1, I}$, $l = \overline{1, L}$, рассчитываются критерии отбраковки K_i , $i = \overline{1, I}$. Если участок не удовлетворяет хотя бы одному из критериев K_i , $i = \overline{1, I}$, то он отбраковывается (участки 2, 5 на рис. 6). На основе (3), с использованием не отбракованных участков (1, 3, 4 на рис. 6), рассчитывается коэффициент корреляции:

$$\rho^* = 0.5 \sum_{j=1}^J (\overline{B}_{1j} - \overline{B}_{0j}) / \sum_{j=1}^J \sqrt{D_{Bj}}. \quad (14)$$

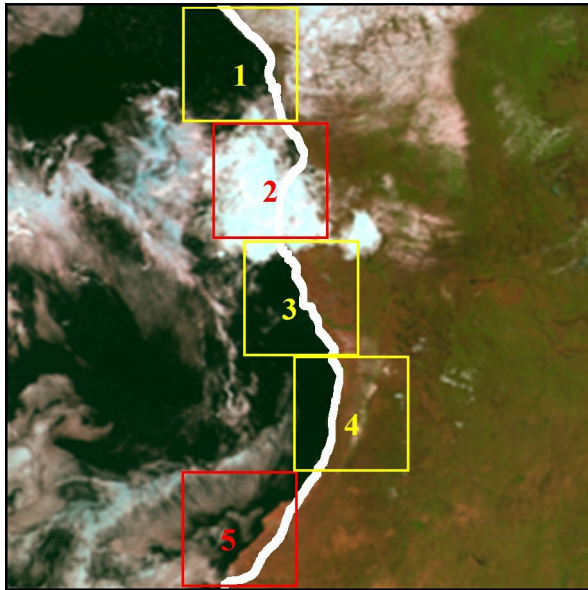


Рис. 6. Устранение участков с облаками (2) и не совпадающими границами (5)

Экспериментально установлено, что эта операция позволяет увеличить до 40% общее число опорных фрагментов и тем самым повысить точность совмещения снимка с картой.

Отбраковка ОТМ, найденных по ЭК. Эта операция основана на сопоставлении статистических характеристик водной поверхности и суши под бинарной маской.

Во-первых, ОТМ отбраковывается, если $B_0 > B_1$ и (или) $D_0 > D_1$.

Во-вторых, ОТМ отбраковывается, если $\rho_{\max} < \rho_0$ и (или) $\rho_{\max} / \rho^* < p$ и (или) $\sigma_\rho < \sigma_\rho^*$, где ρ_{\max} , ρ^* , ρ_0 - соответственно максимальное, следующее по величине в заданной области и пороговое значения коэффициента корреляции ρ ; σ_ρ , σ_ρ^* - соответственно СКО коэффициента корреляции и его пороговое значение.

Отбраковка ОТМ, найденных по контуру диска Земли. Эта операция основана на учете факта пространственной связности и близости контурных точек к некоторой линии, образующей эллипс. Здесь используются три критерия.

Краевые точки отбраковываются, если не удовлетворяют условию

$$R_{\min} - R_{\text{ц}} \leq r \leq R_{\max} + R_{\text{ц}}, \quad (15)$$

где $R_{\text{ц}}$ - радиус круга с центром (0, 0), определяющий допустимую область расположения центра эллипса, $[R_{\min}, R_{\max}]$ - допустимый диапазон изменения радиуса круга, r - расстояние от центра до контурной точки.

Далее исключаются участки контура, заведомо не имеющие отношения к краям диска, при одновременном выполнении условий

$$|\vartheta - \bar{\vartheta}| > 3\sigma_\vartheta, \quad |R - \bar{R}| > 3\sigma_R, \quad (16)$$

где ϑ - направление участка контура, R - радиус кривизны окружности, рассчитанный по окрестности рассматриваемой точки, $\bar{\vartheta}$, \bar{R} и σ_ϑ , σ_R - соответственно средние значения и СКО для ϑ и R .

Наконец, оставшиеся краевые точки бракуются, если

$$|r(\varphi) - \tilde{r}(\varphi)| > \bar{d} + 3\sigma_d, \quad (17)$$

где $r(\varphi)$, $\tilde{r}(\varphi)$ - соответственно эталонный эллипс и функция, построенная по реальным контурным точкам, в полярной системе координат $r\varphi$; \bar{d} и σ_d - среднее значение и СКО, найденные по множеству отчетов разности $r(\varphi) - \tilde{r}(\varphi)$.

Экспериментальные исследования

В ходе экспериментальных исследований получены следующие результаты.

- При нормализации изображений по электронным картам обеспечивается:
 - СКО погрешности порядка 0.2 - 0.5 пикселя;
 - предельная погрешность центральной точки диска Земли 0.7 пикселя;
 - решение поставленной задачи даже при 50%-м затенении изображения.

Эксперименты проводились с привлечением 62-х снимков от спутников «MeteoSat-5» и «MeteoSat-8». В каждом изображении использовалось по 53 ОТМ.

- При нормализации изображений по контурным точкам диска Земли даже при 80 % затенении диска Земли обеспечиваются точность оценивания углов крена и тангажа порядка 3 угл. с и угла рысканья порядка 200 угл. с., что позволило существенно сузить зону поиска ОТМ по электронным картам.

Эксперименты проводились с привлечением 120 изображений от спутника «MeteoSat-8».

- Время нормализации изображений по 53 опорным фрагментам составило 15 секунд, а по контурным точкам диска Земли - 10 секунд. Эксперименты выполнялись на серверной ЭВМ IBM x3850 с 4-мя двудерными процессорами Intel Xeon 7120N 4x3.0GHz / 667 MHz, обладающей оперативной памятью 8 Гбайт.

Литература

1. Злобин В.К., Еремеев В.В. Обработка аэрокосмических изображений. М.: Физматлит, 2006. 288 с.
2. Злобин В.К., Еремеев В.В., Федоткин Д.И. Информационная технология географической привязки космических изображений с использованием электронных карт // Исследование Земли из космоса. 2000. №1. С. 86-91.
3. Rosborough G.W., Baldwin D.G., Emery W.J. Precise AVHRR image navigation // IEEE Trans. Geosci. Rem. Sens. 1994. 32. №3. P.644-657.
4. Mather P.M. Map-image registration accuracy using least-squares polynomials // Int.J. Geogr. Inform. Syst. 1995. 9. №5. P.543-554.

Работа проводилась при финансовой поддержке Федерального агентства по науке и инновациям (госконтракт № 02.740.11.0002).

УДК 004.932: 528.831

ОЦЕНКА РАЗРЕШАЮЩЕЙ СПОСОБНОСТИ АЭРОКОСМИЧЕСКИХ ИЗОБРАЖЕНИЙ НА ОСНОВЕ ИХ СТАТИСТИЧЕСКОГО АНАЛИЗА

Еремеев В.В., д.т.н., директор НИИ «Фотон» РГРТУ

Князьков П.А., аспирант НИИ «Фотон» РГРТУ

Москвитин А.Э., к.т.н., с.н.с. НИИ «Фотон» РГРТУ, foton@rgta.ryazan.ru

Ключевые слова: дистанционное зондирование, пространственное разрешение, аэрокосмические изображения, статистический анализ, функция края, область анализа, ориентация контура, алгоритм, аппроксимация.

Введение

Пространственная разрешающая способность и отношение сигнал/шум являются важнейшими характеристиками систем дистанционного зондирования Земли (ДЗЗ). Они определяют различимость объектов наблюдаемой сцены на фоне шумов. В процессе проектирования систем формирования изображений оценка этих характеристик осуществляется в лабораторных условиях путем подачи на их вход опорных сигналов и специальных мишеней. В процессе эксплуатации систем ДЗЗ выполнить эти операции не представляется возможным, и единственным способом оценки пространственной и радиометрической разрешающей способности является анализ получаемой с помощью систем ДЗЗ видеоинформации.

Известны два подхода к оценке пространственного разрешения систем ДЗЗ по формируемым ими снимкам. Первый подход рассмотрен в работах [1,2]. Он основан на размещении на поверхности Земли тест-объекта в виде шитых белых и черных полос брезента, съемки тест-объекта и анализе степени размытости его границ путем формирования функции рассеивания точки (ФРТ). Второй подход реализован в космическом проекте Quickbird II [3], где на поверхности Земли формируются тест-объекты в виде точечных оптических сигналов. О разрешающей способности судят по степени его расфокусировки на изображении. Общим недостатком этих подходов являются сложность реализации и узкая ориентация на конкретную систему ДЗЗ (на вполне определенный шаг пространственной дискретизации). В связи с этим в настоящей работе рассматривается алгоритм, в котором используется идея первого подхода, но оценка разрешения производится на основе анализа естественных объектов наблюдаемой сцены с контрастными границами.

Рассматривается также технология оценки радиометрического разрешения (отношения сигнал/шум) непосредственно по аэрокосмическим изображениям. В известных работах [4,5] в качестве адекватного описания шума используется аддитивная независимая модель $B = X + \varepsilon$, где B - репродуцируемое изображе-

Представлены две технологии оценки важнейших характеристик систем дистанционного зондирования Земли – пространственной разрешающей способности и отношения сигнал/шум на основе статистического анализа материалов аэрокосмической съемки. Приводятся результаты экспериментальных исследования этих технологий.

ние с шумом ε , X - не зашумленное изображение. Такой моделью описываются электронный шум видеодатчика, шум тракта передачи сигнала и шум квантования [6]. Оценка дисперсии D_ε подобного шума основана на анализе по последовательности отсчетов видеоданных $B_i = X_i + \varepsilon_i$, $i = \overline{1, I}$, отсчетов автокорреляционной функции (АКФ) \widehat{K}_τ , $\tau = 0, 1, 2, \dots$, в которой только нулевой отсчет подвержен искажениям:

$$\widehat{K}_\tau = \begin{cases} K_0 + D_\varepsilon, & \tau = 0; \\ K_\tau, & \tau = 1, 2, \dots, \end{cases} \quad (1)$$

где K_τ , $\tau = 0, 1, 2, \dots$, - не зашумленные отсчеты АКФ.

В настоящей работе рассматривается технология высокоточной оценки D_ε по зашумленному изображению и вопросы оценки точности решения этой задачи.

Оценка пространственной разрешающей способности

Пространственное разрешение систем ДЗЗ наиболее полно описывается через частотно-контрастную характеристику (ЧКХ), которая является модулем Фурье-преобразования функции рассеивания точки (ФРТ) $H(x, y)$. Связь функции спектральной яркости наблюдаемой сцены $E(x, y)$ с непрерывным изображением $B_n(x, y)$ и ФРТ определяется через операцию свертки: $B_n(x, y) = H(x, y) \otimes E(x, y)$, где x и y - горизонтальное и вертикальное направления осей координат. Рассматриваемый ниже алгоритм основан на получении ФРТ путем анализа контрастных границ однородных по яркости объектов изображений по направлениям x и y . Отличительным элементом алгоритма является то, что ФРТ формируется не дифференцированием дискретизованного перехода яркости, в результате чего возникают ошибки, а путем его восстановления с субпиксельной точностью и последующим дифференцированием.

Алгоритм предполагает последовательное выполнение трех этапов: получение по контрастным границам объектов изображения так называемой функции края (ФК); определение по ФК функции рассеивания точки и нахождение по ФРТ искомого ЧКХ. Переходы

$\Phi K \rightarrow \Phi P T \rightarrow \Phi K X$ строго описываются математически, так что центральной задачей является высокоточное нахождение ФК.

Получение ФК. Вначале на изображении определяется прямоугольная область анализа (ОА) $D(m, n)$, $m = \overline{1, M}$ (номер строки), $n = \overline{1, N}$ (номер столбца), которая содержит вертикально ориентированную контрастную границу, пересекающую две противоположные стороны ОА. (рис. 1). Ниже приводимые рассуждения полностью распространяются и для ОА с горизонтальной ориентацией ярких переходов.

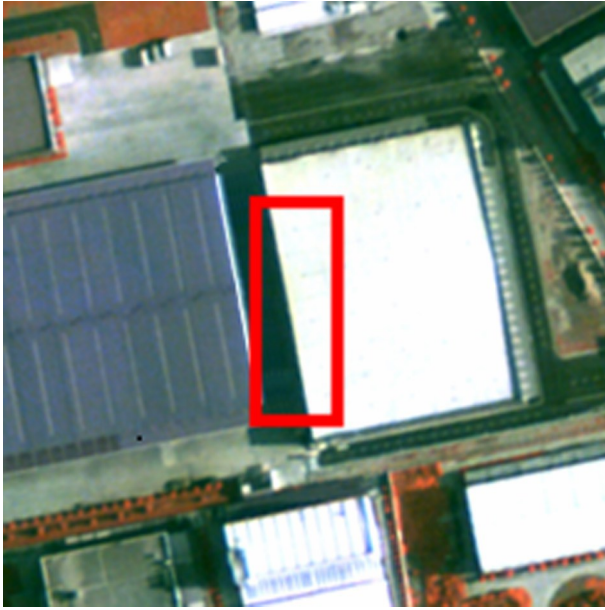


Рис. 1. ОА с вертикальной ориентацией контура

С использованием модуля дискретной производной $P(m, n) = |D(m, n+1) - D(m, n-1)|$, $m = \overline{1, M}$, $n = \overline{2, N-1}$, в ОА выделяются контурные точки $n_m = \arg \max_n P(m, n)$.

Точкам $P(m, n)$, отстоящим по строке от n_m , более чем на допустимый порог размытия контура Δ , присваивается нулевое значение $P(m, n) = 0$. Положение контура, уточненное до долей пикселя, определяется как

$$N_m = \left[\sum_{n=2}^{N-1} n \cdot P(m, n) \right] / \left[\sum_{n=2}^{N-1} P(m, n) \right]. \quad (2)$$

По полученному набору координат (N_m, m) , $m = \overline{1, M}$, по МНК строится геометрическая модель контура

$$x(m) = \sum_{i=0}^S a_i m^i.$$

На основе анализа невязок $N_m - x(m)$ решается вопрос о пригодности ОА для построения ФК.

ФК $g(x)$ формируется как $g(n-x(m)) = D(m, n)$, $m = \overline{1, M}$, $n = \overline{1, N}$, т.е. аргументом этой функции является расстояние точки ОА до контура по оси n , а ее значением яркость в этой точке (рис. 2). Этой операцией и достигается получение ФК с субпиксельной точностью.

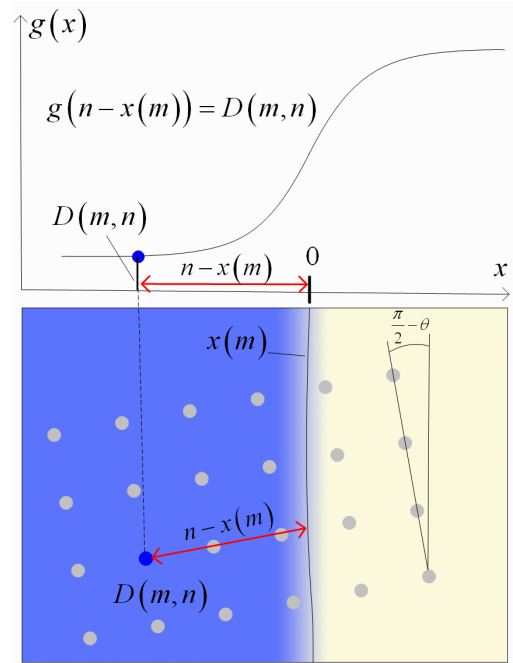


Рис. 2. Принцип формирования ФК

При анализе реальных изображений размеры используемых контуров ограничены, возникает необходимость увеличения статистической выборки для построения ФК. Эта задача решается в предположении, что система ДЗЗ является линейной и инвариантной к сдвигу. В этом случае ФК $g_s(x)$, $s = \overline{1, S}$, полученные по S областям анализа, приводятся к единому нижнему a и верхнему b уровням яркости:

$$\hat{g}_s(x) = \left(\frac{b-a}{b_s-a_s} \right) (g_s(x) - a_s) + a, \quad s = \overline{1, S}, \quad (3)$$

где $\hat{g}_s(x)$ - приведенные к единому яркостному диапазону ФК, a_s и b_s - нижние и верхние уровни $g_s(x)$.

По мере отклонения угла наклона яркостного перепада θ от вертикальной оси увеличивается масштаб ФК μ по горизонтали. Из рис. 3 следует, что $\mu = R/R_\theta = 1/\sin \theta$, где R - радиус кривизны ФРТ в окрестности точки ее встречи с яркостным перепадом.

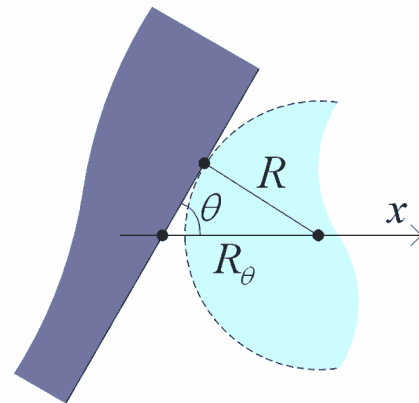


Рис. 3. Учет наклона контура

Поэтому функции края $\hat{g}_s(x)$, $s = \overline{1, S}$, полученные при различных углах θ_s , необходимо привести к $\theta = \pi/2$, сжав их по оси x в $1/\mu_s = \sin \theta_s$ раз относи-

тельно точки $x = 0$. После масштабирования функций края по двум направлениям операция их объединения становится корректной. В результате формируется совокупная функция края

$$g^*(x) = \bigcup_{s=1}^S \widehat{g}_s(x/\mu_s),$$

которая содержит измерения от всех областей анализа.

Определение ФРТ и ЧКХ. Точной оценке ФРТ по $g^*(x)$ мешают три фактора: наличие на изображении шумов; неоднородность по яркости объекта и фона; неточность описания контура $x(m)$. Для уменьшения влияния этих факторов исследованы два вида аппроксимаций ФК - кусочно-полиномиальная и параметрическая.

При кусочно-полиномиальной аппроксимации для точек, равномерно расположенных на оси x , в скользящей вдоль этой оси апертуре по МНК строятся полиномы с привлечение всех точек, попавших в апертуру. Уточненное значение отсчета ФК в центральной точке апертуры определяется через нулевой коэффициент полинома. Разности соседних уточненных отсчетов ФК определяют отсчеты ФРТ.

При параметрической аппроксимации ФК представляется аналитическими моделями вида

$$q_j(x, k) = 0.5(a+b) + |a-b|\varphi_j(x, k), \quad j=1, 2, 3, \quad (4)$$

где $\varphi_1(x, k) = \frac{1}{\pi} \arctg(kx)$,

$$\varphi_2(x, k) = 0.5 \left(\frac{e^{kx} - e^{-kx}}{e^{kx} + e^{-kx}} \right), \quad \varphi_3(x, k) = \frac{1}{k\sqrt{2\pi}} \int_0^x e^{-2k^2 y^2} dy.$$

Параметр k и функция q_j определяются из условий:

$$e_j = \sum_x (g^*(x) - q_j(x, k))^2 = \min,$$

$$j = \arg(\min(e_1, e_2, e_3)).$$

Искомая ФРТ находится как производная выбранной q_j , а ЧКХ как нормированный модуль спектра ФРТ.

Оценка уровня шума

В соответствии с (1) оценка дисперсии шума основана на экстраполяции значения нулевого отсчета АКФ по отсчетам при $\tau = 1, 2, 3, \dots$. Для решения задачи в такой постановке необходимо выбрать модель АКФ, которая была бы адекватной и наиболее точно предсказывала искомый нулевой отсчет.

Модели автокорреляционной функции. Исходя из свойств АКФ реальных изображений, ее аналитическая модель K_τ должна удовлетворять следующим свойствам:

- 1) быть четной функцией относительно τ ($K_{-\tau} = K_\tau$);
- 2) иметь максимум в точке $\tau = 0$, ($K_0 \geq K_\tau, \tau = 1, 2, 3, \dots$);
- 3) асимптотически стремиться к нулю при увеличении τ ($K_\tau > K_{\tau+1}$ при $\tau = 0, 1, 2, \dots$ и $\lim_{\tau \rightarrow \infty} K_\tau = 0$).

С учетом этого исследованы две двухпараметрические модели АКФ:

$$K_{1\tau} = a_1 e^{-c_1 \tau^2}, \quad K_{2\tau} = \frac{a_2}{1 + c_2 \tau^2}. \quad (5)$$

Параметры моделей определяются по отчетам АКФ при $\tau = 1, 2$ как

$$a_1 = \sqrt[3]{\widehat{K}_1^4 / \widehat{K}_2}, \quad c_1 = \ln \sqrt[3]{\widehat{K}_1 / \widehat{K}_2},$$

$$a_2 = 3\widehat{K}_1 \widehat{K}_2 / (4\widehat{K}_2 - \widehat{K}_1), \quad c_2 = (\widehat{K}_1 - \widehat{K}_2) / (4\widehat{K}_2 - \widehat{K}_1), \quad (6)$$

Из (5) и (6) следуют оценки дисперсии шума соответственно для моделей $K_{1\tau}, K_{2\tau}$:

$$D_{1\epsilon} = \widehat{K}_0 - \sqrt[3]{\widehat{K}_1^4 / \widehat{K}_2}, \quad D_{2\epsilon} = \widehat{K}_0 - 3\widehat{K}_2 \widehat{K}_1 / (4\widehat{K}_2 - \widehat{K}_1). \quad (7)$$

Экспериментальные исследования, проведенные с привлечением изображений от различных сканирующих устройств, показали, что удовлетворительная точность оценки D_ϵ обеспечивается лишь при достаточно большом уровне шума. Это связано с тем, что на практике $D_\epsilon \ll K_\epsilon, \tau = 0, 1, 2, \dots$. Поэтому малые ошибки в определении K_0 приводят к значительным относительным погрешностям оценки D_ϵ . Рассмотрим подход, свободный от этого недостатка. Он основан на корреляционном анализе разностного изображения, для которого значения АКФ сопоставимы по величине с D_ϵ .

Оценка D_ϵ по разностному изображению. По исходным отсчетам яркости B_i , формируется последовательность разностей:

$$Z_i = B_i - B_{i-1} = (X_i - X_{i-1}) + (\epsilon_i - \epsilon_{i-1}). \quad (8)$$

Отсчеты АКФ \widehat{R}_τ , полученные по последовательности Z_i , равны:

$$\widehat{R}_0 = R_0 + 2D_\epsilon, \quad \widehat{R}_1 = R_1 - D_\epsilon, \quad \widehat{R}_2 = R_2, \quad \widehat{R}_3 = R_3, \quad (9)$$

где $R_\tau = -(K_{\tau-1} - 2K_\tau + K_{\tau+1})$.

Следовательно, значения R_τ численно равны второй дискретной производной от K_τ с обратным знаком. Поэтому в качестве моделей поведения R_τ логично использовать вторые производные функций $K_{1\tau}, K_{2\tau}$ (5) с противоположным знаком, которые вблизи точки $\tau = 0$ хорошо описывают корреляционные свойства разностного изображения:

$$R_{1\tau} = -K_{1\tau}'' = 2ace^{-c\tau^2} (1 - 2c\tau^2), \quad (10)$$

$$R_{2\tau} = -K_{2\tau}'' = 2ac \frac{(1 - 3c\tau^2)}{(1 + c\tau^2)^3}. \quad (11)$$

С учетом (9) для модели $R_{1\tau}$ дисперсия шума D_{R1} определяется из системы

$$\widehat{R}_{10} = 2ac + 2D_{R1},$$

$$\widehat{R}_{11} = 2ace^{-c} (1 - 2c) - D_{R1},$$

$$\widehat{R}_{12} = 2ace^{-4c} (1 - 8c) \quad (12)$$

и равна

$$D_{R1} = \frac{1}{2} \widehat{R}_{10} - \frac{\widehat{R}_{12} e^{4c}}{2(1 - 8c)}. \quad (13)$$

Для модели $R_{2\tau}$ дисперсия шума D_{R2} определяется из системы

$$\hat{R}_{20} = 2ac + 2D_{R2},$$

$$\hat{R}_{21} = \frac{2ac(1-3c)}{(1+c)^3} - D_{R2},$$

$$\hat{R}_{22} = \frac{2ac(1-12c)}{(1+4c)^3}.$$

$$\text{Она равна } D_{R2} = \frac{1}{2} \left(\hat{R}_{20} - \hat{R}_{22} \frac{(1+4c)^3}{1-12c} \right). \quad (14)$$

На основе полученных оценок дисперсии шума не представляет труда определить оценку отношения сигнал/шум, например, как $\gamma = \sqrt{(D_B - D_\epsilon)/D_\epsilon}$, либо как отношение заданного уровня сигнал к $\sqrt{D_\epsilon}$. Параметр γ определяет радиометрическое разрешение системы ДЗЗ.

Экспериментальные исследования

Рассмотренные в настоящей работе алгоритмы оценки пространственного и яркостного разрешения на основе статистического анализа изображений экспериментально исследованы с привлечением реальных космических снимков земной поверхности от различных систем ДЗЗ.

Оценка пространственного разрешения. Схема экспериментальных исследований точности оценки пространственного разрешения представлена на рис. 4.

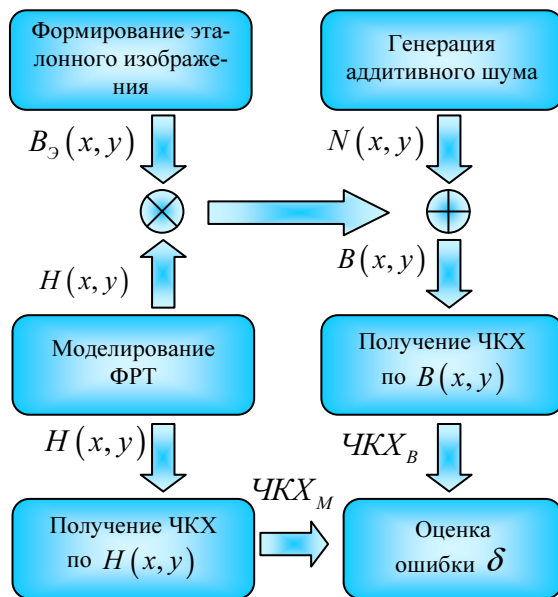


Рис. 4. Схема эксперимента

В начале формируется эталонное изображение $B_3(x, y)$, содержащее идеальный яркостной перепад. В это изображение вносятся вполне определенные искажения посредством свертки $B_3(x, y)$ и моделируемой ФРТ $H(x, y)$ и добавления аддитивного шума $N(x, y)$ с СКО σ_N . В результате формируется изображение $B(x, y)$. Путем сопоставления $ЧКХ_B$, полученной по $B(x, y)$ рассмотренным выше алгоритмом, с модельной $ЧКХ_M$ определяется относительная погрешность оценки пространственного разрешения δ . В качестве модельной ФРТ использована гауссоида

$$H(x, y) = \frac{1}{\sqrt{2\pi}a} e^{-\frac{x^2+y^2}{2a^2}},$$

где a - СКО, определяющее степень размытия границ [7].

Результаты экспериментов представлены в таблице.

Таблица
Ошибка оценки разрешения δ

№ эксперимента	a	σ_N	$\delta, \%$	
			Кусочно-полином.	Параметрическая
1	0,5	1	6,2	6,1
2	0,5	1,5	10,5	8,6
3	0,7	1	6,1	6,1
4	0,7	1,5	8,5	7,2

Из таблицы следует, что кусочно-полиномиальная и параметрическая аппроксимация ФК дают близкие результаты. Однако, кусочно-полиномиальная модель дает возможность получения ЧКХ для систем с произвольной ФРТ, но является более чувствительной к шуму изображения; напротив, параметрическая модель более устойчива к шуму, но ее возможно применять только для систем, ФРТ которых могут быть достоверно представлены производной от функций q_j .

Оценка уровня шума. В результате экспериментальных исследований двух рассмотренных выше технологий оценки дисперсии шума (по исходному и разностному изображению) установлено:

- СКО оценки дисперсии шума по исходному изображению составляет для каждой из моделей АКФ примерно одинаковую величину $\sigma_D \approx 0,6$ градации яркости. Следовательно, технологию оценки уровня шума по исходному изображению можно использовать только при высоком уровне шумов, порядка $D_\epsilon \approx 5-7$ и выше. В этом случае относительная ошибка оценки D_ϵ , как правило, не будет превышать 10%.

- СКО оценки дисперсии шума по разностному изображению для модели $R_{1\tau}$ составляет порядка $\sigma_D \approx 0,25$, а для модели $R_{1\tau}$ порядка $\sigma_D \approx 0,35$ градации яркости. Следовательно, технологию определения оценки уровня шума по разностному изображению можно использовать для оценки как высоких уровней шума, так и малых уровней шумов, порядка $D_\epsilon \approx 2-3$ и выше.

Литература

1. Stanley A. Morain, Amelia M. Budge. Post-Launch Calibration of Satellite Sensors. Proceeding of the international workshop on radiometric and geometric calibration, 2-5 December 2003. Earth Data Analysis Center, University of New Mexico, Albuquerque, NM, USA.
2. Brian M. Bleeze. Modeling the MTF and Noise Characteristics of Complex Image Formation Systems // Center for Science Rochester Institute of Technology, May 15, 1998.
3. Manjunath Kempaiah Rangaswamy. Quickbird II Two-dimensional On-orbit Modulation Transfer Function Analysis Using Convex Mirror Array. South Dakota State University, 2003.
4. Ярославский Л.П. Введение в цифровую обработку изображений. М.: Сов. радио, 1979. – 312 с..
5. Liu C., Freeman W., Szeliski R., Kang S. Noise estimation from a single image // IEEE Conf. Computer Vision and Patter Recognition, 2001. – P. 901–908.
6. Schreiber W. Fundamentals of electronic imaging systems. Springer-Verlag, 1986. – 268 p.
7. Brian L. Markham. The Landsat Sensors' Spatial Responses // IEEE Transactions on Geoscience and Remote Sensing, Vol. GE-23, No. 6, November 1985.

УДК 004.932.2

РАДИОМЕТРИЧЕСКАЯ НОРМАЛИЗАЦИЯ ИЗОБРАЖЕНИЙ ОТ МНОГОМАТРИЧНЫХ СКАНИРУЮЩИХ УСТРОЙСТВ

Еремеев В.В., д.т.н., директор НИИ «Фотон» РГРТУ

Зенин В.А., аспирант НИИ «Фотон» РГРТУ, gislab@org.etr.ru

Ключевые слова: дистанционное зондирование Земли, радиометрическая нормализация, скан, прибор с зарядовой связью, функция распределения, автоковариационная функция, средняя яркость.

Введение

В практике дистанционного зондирования Земли широкое применение нашли многозональные сканирующие устройства, построенные на основе приборов с зарядовой связью: ПЗС-линеек и ПЗС-матриц. Несмотря на высокие технические характеристики таких приборов, им присущ серьезный недостаток. При использовании ПЗС-линеек из-за различий передаточных характеристик (чувствительностей и темновых сигналов) отдельных фотоприемников на изображении появляются искажения в виде характерной «полосатости». Изображение от многоматричного видеодатчика формируется в виде вертикальных полос (сканов) с их небольшим перекрытием в горизонтальном (строчном) направлении. Скан включает набор видеоданных от вполне определенной ПЗС-матрицы. В виду неодинаковых передаточных характеристик ПЗС-матриц и различий условий съемки изображение представляет собой резко отличающиеся по яркости полосы (рис. 1).

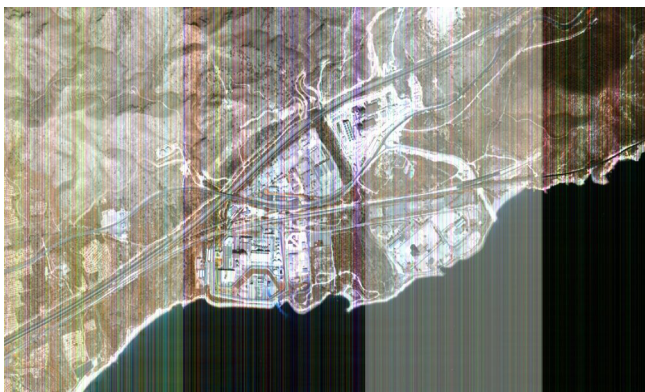


Рис. 1. Фрагмент изображения от многоматричного видеодатчика

Коррекцию подобных искажений можно выполнить с помощью двух технологий [1]: по данным бортовой калибровки датчика и путем статистического анализа изображения, искаженного структурным шумом. По данным бортовой калибровки высокое качество коррекции изображения может быть достигнуто только при высокостабильных калибровочных сигналах и их неизменности по

Представлено решение задачи статистической радиометрической коррекции структурных искажений на спутниковых изображениях земной поверхности, формируемых с помощью многоматричных видеодатчиков. Рассмотрены алгоритмы коррекции изображений, содержащих существенно неоднородные по яркости объекты, когда известные алгоритмы оказываются не работоспособными. Предложены подходы по яркостному выравниванию видеоданных от различных фотоприемных матриц и устранение структурного шума, обусловленного различием чувствительности и темновых сигналов отдельных фотоприемных элементов. Приведены результаты обработки реальных снимков.

всему полю зрения датчика, что трудно реализуемо на практике. Статистические подходы не требуют наличия сложной калибровочной аппаратуры и позволяют решить поставленную задачу с привлечением собственно искаженного изображения, поэтому в последнее время делаются активные попытки их использования.

Один из статистических подходов основан на приведении эмпирической функций распределения яркости корректируемого k -го столбца $F_k(b)$, $k=1, \overline{K}$, к эталонной функции $F_{k_0}(b^*)$, которая формируется по симметричной апертуре столбцов: $F_k(b) = F_{k_0}(b^*)$ [2], где b и b^* - исходное и скорректированное значения яркости. Отсюда функция коррекции $\varphi_k(b)$ определяется как

$$b^* = \varphi_k(b) = F_{k_0}^{-1}(F_k(b)), \quad k=1, \overline{K}. \quad (1)$$

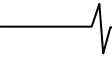
В принципе, такой подход может быть использован и для яркостного выравнивания сканов. В этом случае видеоданные одного из сканов (например правого), находящиеся в зоне перекрытия, используются для формирования $F_k(b)$, а видеоданные соседнего скана, взятые из той же зоны перекрытия, используются для формирования $F_{k_0}(b^*)$.

Линейная модель коррекции

Установлено, что для ПЗС функция $\varphi_k(b)$ в (1) линейна: $b^* = g_k b + a_k$. В ней мультипликативный коэффициент g_k характеризует чувствительность фотоприемника, а аддитивный a_k - общее смещение яркости (темновой сигнал). Находя средние яркости $\overline{b_k^*} = g_k \overline{b_k} + a_k$ и СКО $\overline{\sigma_k^*} = g_k \overline{\sigma_k}$ по всем элементам столбца и апертуры, определяются искомые коэффициенты функций коррекции структурного шума:

$$g_k = \overline{\sigma_k^*} / \overline{\sigma_k}, \quad a_k = \overline{b_k^*} - \overline{b_k} g_k, \quad k=1, \overline{K} \quad (2)$$

Такая модель дает хорошие результаты, если изображение не содержит дополнительного аддитивного шума, что на практике случается очень редко. Наложение



ние на сигнал b_k аддитивного некоррелированного шума ε_k ($\tilde{b}_k = b_k + \varepsilon_k$) приводит к увеличению дисперсии $\tilde{\sigma}_k^2 = \sigma_k^2 + \sigma_\varepsilon^2$ и искажению коэффициентов коррекции: $\tilde{g}_k = \sigma_k^* / \sqrt{\sigma_k^2 + \sigma_\varepsilon^2}$, $\tilde{a}_k = b_k^* - b_k g_k$. В результате влияние шума приведет к появлению остаточных структурных искажений.

Для устранения такой методической ошибки предлагается при нахождении мультипликативного коэффициента использовать не СКО (дисперсию), а автоковариационную функцию (АКФ) в столбцовом направлении, любой отсчет которой, кроме нулевого, $\mu_{k\tau}$, $\tau = 1, 2, 3, \dots$, не зависит от уровня шума. Действительно, для зашумленного сигнала

$$\mu_{k\tau} = \frac{1}{N-\tau} \sum_{j=1}^{N-\tau} b_j b_{j+\tau} - \frac{1}{(N-\tau)^2} \sum_{j=1}^{N-\tau} b_j \sum_{j=1}^{N-\tau} b_{j+\tau},$$

а АКФ зашумленного сигнала

$$\begin{aligned} \tilde{\mu}_{k\tau} &= \frac{1}{N-\tau} \sum_{j=1}^{N-\tau} (b_j + \varepsilon_j)(b_{j+\tau} + \varepsilon_{j+\tau}) - \\ &= \frac{1}{(N-\tau)^2} \sum_{j=1}^{N-\tau} (b_j + \varepsilon_j) \sum_{j=1}^{N-\tau} (b_{j+\tau} + \varepsilon_{j+\tau}) = \\ &= \frac{1}{N-\tau} \sum_{j=1}^{N-\tau} (b_j b_{j+\tau} + b_j \varepsilon_{j+\tau} + \varepsilon_j b_{j+\tau} + \varepsilon_j \varepsilon_{j+\tau}) - \\ &= \frac{1}{(N-\tau)^2} \sum_{j=1}^{N-\tau} b_j \sum_{j=1}^{N-\tau} b_{j+\tau} + \sum_{j=1}^{N-\tau} \varepsilon_j \sum_{j=1}^{N-\tau} b_{j+\tau} + \sum_{j=1}^{N-\tau} b_j \sum_{j=1}^{N-\tau} \varepsilon_{j+\tau} + \sum_{j=1}^{N-\tau} \varepsilon_j \sum_{j=1}^{N-\tau} \varepsilon_{j+\tau} = \\ &= \mu_{k\tau} + 2\text{cov}(b, \varepsilon) + \text{cov}(\varepsilon, \varepsilon) = \mu_{k\tau}. \end{aligned}$$

Поэтому вместо (2) определять коэффициенты коррекции структурного шума следует через АКФ:

$$g_k = \sqrt{\mu_{k\tau}^* / \mu_{k\tau}}, \quad a_k = b_k^* - b_k \sqrt{\mu_{k\tau}^* / \mu_{k\tau}}, \quad k = \overline{1, K}.$$

Линейная модель может быть эффективно использована для яркостного выравнивания сканов, поскольку в небольших зонах их перекрытия содержатся данные от одной и той же сцены. В то же время при коррекции структурных искажений, обусловленных неодинаковыми чувствительностями и темновыми сигналами отдельных фотоприемников, эта модель дает удовлетворительный результат лишь при статистически однородной сцене. В случаях, когда наблюдаемый сюжет содержит резко отличающиеся по яркости объекты, эталонное значение АКФ $\mu_{k\tau}^*$, полученное по апертуре столбцов, может существенно отличаться от $\mu_{k\tau}$, найденного для корректируемого столбца, и эти отличия могут быть обусловлены не столько изменением характеристик фотоприемников, а главным образом, изменением сюжета в пределах апертуры. В этом случае необходимо вначале выполнить яркостное выравнивание сканов, а затем коррекцию данных, относящихся к отдельным фотоприемникам, приняв специальные меры по нейтрализации влияния неоднородности сюжета.

Яркостное выравнивание сканов

Пусть изображение включает I сканов, которые пронумеруем слева направо $i = \overline{1, I}$. Для двух соприкасающихся сканов с номерами $i-1$ и i определим наборы X_{i-1} и X_i данных в зоне перекрытия (рис. 2).

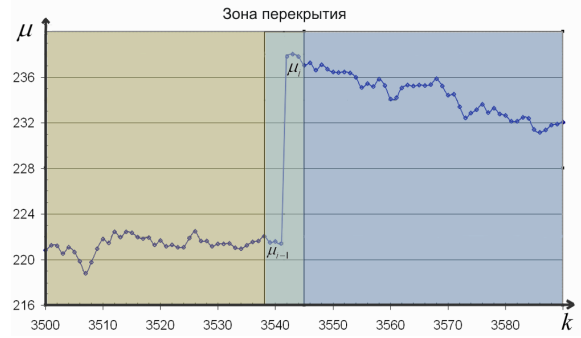


Рис. 2. Изменение АКФ на границе сканов

Сначала получим относительные параметры линейных преобразований, принимая левые зоны перекрытия как базовые, а правые как корректируемые:

$$\hat{r}_i = \mu_{i-1} / \mu_i, \quad \hat{c}_i = \bar{X}_{i-1} - \hat{r}_i \bar{X}_i, \quad i = \overline{2, I}, \quad (3)$$

где μ_{i-1} , \bar{X}_{i-1} и μ_i , \bar{X}_i - значения АКФ и средней яркости, полученные по данным зоны перекрытия для $(i-1)$ -го и i -го сканов; r_i , c_i - мультипликативный и аддитивный коэффициенты линейного преобразования данных i -го скана к $(i-1)$ -му.

Затем найдем при заданных параметрах преобразования первой полосы r_1 и c_1 коэффициенты r_i и c_i , $i = \overline{2, I}$, абсолютного преобразования всех других полос, при которых достигается яркостное выравнивание всего изображения:

$$\begin{aligned} r_1 \bar{X}_1 + c_1 &= \hat{r}_2 (\hat{r}_1 \bar{X}_1 + c_1) + \hat{c}_2 = \\ &= \hat{r}_3 \hat{r}_2 (\hat{r}_1 \bar{X}_1 + c_1) + \hat{c}_2 + \hat{c}_3. \end{aligned}$$

Отсюда следуют рекуррентные соотношения:

$$r_i = r_{i-1} \hat{r}_i, \quad c_i = c_{i-1} \hat{r}_i + \hat{c}_i, \quad i = \overline{2, I}. \quad (4)$$

Остается открытым вопрос выбора r_1 и c_1 . Рассмотрим два возможных подхода.

1-й подход. Зададим r_1 и c_1 такими, чтобы средние яркости и средние дисперсии изображений до и после преобразований совпадали, т.е. выполнялись бы два условия:

$$\sum_{i=1}^I \bar{B}_i = \sum_{i=1}^I (r_i \bar{B}_i + c_i), \quad (5)$$

$$\sum_{i=1}^I D_i = \sum_{i=1}^I r_i^2 D_i, \quad (6)$$

где \bar{B}_i и D_i - средняя яркость и дисперсия для i -го скана.

Из (6) с учетом (4) следует:

$$\sum_{i=1}^I D_i = r_1^2 D_1 + r_1^2 \sum_{i=2}^I P_i^2 D_i, \quad (7)$$

где $P_i = \prod_{j=2}^i \hat{r}_j$, $P_i = P_{i-1} \hat{r}_i$, $P_1 = 1$, $i = \overline{2, I}$.

Из (7) получим

$$r_1 = \sqrt{\left(\sum_{i=1}^I D_i \right) / \left(D_1 + \sum_{i=2}^I P_i^2 D_i \right)}. \quad (8)$$

Параметр c_1 находится из условия (5) и второго соотношения в (4):

$$\sum_{i=1}^I \bar{B}_i = \sum_{i=1}^I r_i \bar{B}_i + \sum_{i=1}^I c_i.$$

Выделим из последней суммы члены с c_1 :

$$\sum_{i=1}^I c_i = c_1 + c_1 \sum_{i=2}^I P_i + \sum_{i=2}^I S_i,$$

где $S_i = S_{i-1} \hat{r}_i + \hat{c}_i$, $S_1 = 0$, $i = \overline{2, I}$.

С учетом этого получим

$$\sum_{i=1}^I (1-r_i) \bar{B}_i = c_1 \left(1 + \sum_{i=2}^I P_i \right) + \sum_{i=2}^I S_i,$$

$$c_1 = \left(\sum_{i=1}^I (1-r_i) \bar{B}_i - \sum_{i=2}^I S_i \right) / \left(1 + \sum_{i=2}^I P_i \right). \quad (9)$$

2-й подход. Зададим вначале $r_1 = 1$, $c_1 = 0$; получим по (4) коэффициенты r_i , c_i $i = \overline{2, I}$; и для всего скорректированного изображения B_i^* найдем такие r и c , при которых

$$\sum_{i=1}^I \bar{B}_i = \sum_{i=1}^I (r \bar{B}_i^* + c), \quad \sum_{i=1}^I D_i = \sum_{i=1}^I r^2 D_i^*, \quad (10)$$

где \bar{B}_i^* , D_i^* - средняя яркость и дисперсия изображения B_i^* .

Из (10) найдем:

$$\sum_{i=1}^I \bar{B}_i = r \sum_{i=1}^I \bar{B}_i^* + I c, \quad \sum_{i=1}^I D_i = r^2 \sum_{i=1}^I D_i^*;$$

$$r = \sqrt{\left(\sum_{i=1}^I D_i \right) / \sum_{i=1}^I D_i^*}, \quad c = \frac{1}{I} \sum_{i=1}^I \bar{B}_i - \frac{r}{I} \sum_{i=1}^I \bar{B}_i^*, \quad (11)$$

где $D_i^* = r_i^2 D_i$, $B_i^* = r_i B_i + c_i$, ($r_1 = 1$, $c_1 = 0$); B_i , D_i - исходное изображение и его дисперсия.

Окончательные значения коэффициентов, при которых выполняются условия (10), равны:

$$r_1^* = r, \quad c_1^* = c, \quad r_i^* = r_i r, \quad c_i = c + r c_i, \quad i = \overline{2, I}.$$

В таблице представлен пример корректирующих коэффициентов для 24-х сканов изображения от КА «Ресурс-ДК».

Яркостное выравнивание отдельных столбцов

Рассмотрим несколько подходов к нахождению функций коррекции структурных искажений для случая, когда наблюдаемая сцена содержит резко отличающиеся по яркости объекты, чаще всего это водная поверхность, суша и облака.

Алгоритм на основе нормирования функций распределения. Изображениям, содержащим существенно неоднородные по яркости объекты, соответствуют гистограммы $g_k(b)$, $k = \overline{1, K}$, имеющие ярко выраженную многомодальную форму, обычно в виде двух – трех «горбов», разделенных протяженными «впадинами» с $g_k(b) \approx 0$. Из-за изменения сюжета эти гистограммы могут отличаться как по числу «горбов», так и по представительству (объему) данных в одноименных «горбах». В этих условиях известные статистические алгоритмы, основанные на оценке чувствительности и темнового сигнала для каждого ПЗС-элемента путем сопоставления эмпирических функций распределения $F_k(b)$, $k = \overline{1, K}$, оказываются неработоспособными. Рассмотрим алгоритм коррекции структурных искажений, основанный на идентификации и нормировании данных, соответствующих одноименным «горбам».

Таблица.

Корректирующие коэффициенты

Номер скана	№ столбца		r_i	c_i
	начало	конец		
1	1	232	1,016	-5,377
2	233	505	0,992	-1,641
3	506	778	1,002	-1,612
4	779	1010	0,997	-1,286
5	1013	1244	1,001	0,728
6	1245	1517	0,992	0,528
7	1518	1790	1,03	14,603
8	1791	2022	1,028	15,98
9	2025	2256	1,028	-0,927
10	2257	2529	1,04	-1,324
11	2530	2802	1,042	-2,277
12	2803	3034	1,031	-1,405
13	3037	3268	1,02	0,094
14	3269	3541	1,012	0,685
15	3542	3814	0,976	-6,089
16	3815	4046	1,003	-8,474
17	4049	4280	0,995	0,227
18	4281	4553	0,991	1,897
19	4554	4826	0,986	1,475
20	4827	5058	0,98	3,49
21	5061	5292	0,946	-3,389
22	5293	5565	0,936	-0,753
23	5566	5838	0,912	-2,233
24	5839	6070	0,92	-2,921

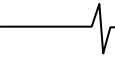
Вначале формируются $F_k(b)$, $k = \overline{1, K}$, и в табличной форме находятся обратные функции $F_{k_p} \rightarrow b_{k_p}$, где $F_{k_p} = p/(P+1)$, $p = \overline{1, P}$. Затем фиксируется принадлежность значения яркости b_{k_p} , $p = \overline{1, P}$, «горбу» ($\Delta_{k_p} \leq \Delta_{k,med} + \Delta_k$) или «впадине» ($\Delta_{k_p} \geq \Delta_{k,med} + \alpha \Delta_k$), где $\Delta_{k_p} = b_{k_p} - b_{k,p-1}$, $\Delta_{k,med} = \text{med } \Delta_{k_p}$, $\Delta_{k,min} = \min \Delta_{k_p}$, $\Delta_k = \Delta_{k,med} - \Delta_{k,min}$, α – коэффициент, характеризующий протяженность «впадины» ($\alpha = 3-5$). После этого выполняется коррекция функции распределения $F_k(b) \rightarrow F_k^*(b)$ так, чтобы все $F_k^*(b)$, $k = \overline{1, K}$, в центрах «впадин» b_{k_m} (m – номер «впадины») имели бы одинаковые значения:

$$F_k^*(b) = \hat{F}_{m-1} + \frac{\hat{F}_m - \hat{F}_{m-1}}{\hat{F}_{k_m} - \hat{F}_{k,m-1}} (F_k(b) - \hat{F}_{k,m-1}), \quad (12)$$

$$\hat{b}_{k,m-1} \leq b \leq \hat{b}_{k_m}, \quad m = \overline{1, M}$$

где $\hat{F}_{k_m} = F_k(\hat{b}_{k_m})$, $\hat{F}_m = \frac{1}{K} \hat{F}_{k_m}$, $\hat{F}_0 = 0$.

Преобразования (12) обеспечивают равное представительство данных в одноименных «горбах» для всех $F_k^*(b)$, тем самым нейтрализуют изменение сюжета в строчном направлении.



Далее формируется таблица $F_k^* \rightarrow b_{kp}^*$, $p = \overline{1, P}$, на основе которой определяются корректирующие функции: вначале в табличной форме $b_{kp}^* \rightarrow y_{kp}^*$,

$$\text{где } y_{kp}^* = \frac{1}{P} \sum_p b_{kp}^*,$$

а затем в виде полинома $y = P_k(b)$, $k = \overline{1, K}$.

Такой алгоритм позволяет выполнить коррекцию структурного шума на существенно неоднородных по яркости изображениях. Однако при слабой различимости «горбов» на гистограмме могут возникать ошибки отнесения видеоданных к тому или иному «горбу», что приводит к потере качества коррекции.

Алгоритм на основе кластеризации объектов изображения. Более качественные результаты дают подходы, основанные на разделении изображения на области с близкими характеристиками, т.е. на использовании процедуры кластеризации. Рассмотрим один из таких подходов [3]. Сначала выполняется предварительная коррекция структурных искажений с использованием линейной модели. Далее элементы изображения разделяются по признаку принадлежности суши или водной поверхности. Порог кластеризации Z_b определяется по гистограмме $g(b)$ предварительно откорректированного изображения как точка максимума функции межклассовой дисперсии

$$\sigma_B^2(b_0) = p_0(b_0)(\gamma_0(b_0) - \gamma)^2 + p_1(b_0)(\gamma_1(b_0) - \gamma)^2, \quad (13)$$

где b_0 - пробный порог; $p_0(b_0) = \sum_{b=0}^{b_0} g(b)$,

$$p_1(b_0) = \sum_{b=b_0+1}^{B_{\max}} g(b) - \text{соответственно оценки вероятностей}$$

попадания произвольного пикселя в класс водной поверхности или суши; $\gamma_0(b_0) = \sum_{b=0}^{b_0} b g(b) / p_0(b_0)$,

$$\gamma_1(b_0) = \sum_{b=b_0+1}^{B_{\max}} b g(b) / p_1(b_0) \text{ и } \gamma - \text{соответственно}$$

нормированные средние яркости в пределах классов водной поверхности, суши и изображения в целом.

Можно показать, что

$$Z_b = \arg \max_{b_0} \frac{[\gamma p_0(b_0) - \gamma(b_0)]^2}{p_0(b_0)[1 - p_0(b_0)]}, \quad (14)$$

где $\gamma(b_0) = \sum_{b=0}^{b_0} b g(b)$ - не нормированная средняя яркость изображения.

Для повышения надежности кластеризации используется модуль дискретной производной яркости в столбцовом направлении, который инвариантен к изменению темновых сигналов фотоприемников. В этом случае порог кластеризации Z_Δ определяется аналогичным образом по гистограмме модуля дискретных производных. Элемент изображения относится к водной поверхности при $b_{kn} \leq Z_c$ и $b_{kn} - b_{k,n-1} \leq Z_\Delta$, где n - номер строки изображения.

После разделения изображения на области водной поверхности и суши, коррекция структурных искажений

выполняется для каждого кластера в отдельности по одному из рассмотренных выше алгоритмов.

Этот алгоритм имеет тот же недостаток, что и предыдущий: при перекрытии яркостных диапазонов водной поверхности и суши кластеризация изображения выполняется с потерей качества.

Алгоритм на основе фильтрации неоднородного сюжета. Рассмотрим некоторый корректируемый столбец изображения B_k . Образует относительно него симметричную апертуру из $(2S+1)$ столбцов. Разобьем эту апертуру в кадровом направлении на равновеликие фрагменты B_{vk} , $v = \overline{1, V}$, включающие по N строк. В пределах фрагмента для каждого столбца, $s = -S, S$, вычислим значения АКФ:

$$\mu_{v,k+s} = \frac{1}{N-1} \sum_{j=1}^{N-1} b_{v,k+s,j} b_{v,k+s,j+1} - \frac{1}{(N-1)^2} \sum_{j=1}^{N-1} b_{v,k+s,j} \sum_{j=1}^{N-1} b_{v,k+s,j+1},$$

$$s = \overline{-S, S}, v = \overline{1, V}. \quad (15)$$

Эталонное значение АКФ для v -го фрагмента определим как

$$\hat{\mu}_{v,k} = \text{med}_s(\mu_{v,k+s}, s = \overline{-S, S}). \quad (16)$$

Выбор в качестве эталона АКФ фрагмента медианного значения позволяет отфильтровать данные по тем столбцам фрагмента, сюжет которых резко отличается от сюжета большинства столбцов. На рис. 3 приведены профили АКФ столбцов и соответствующих эталонов, построенные по реальным данным.

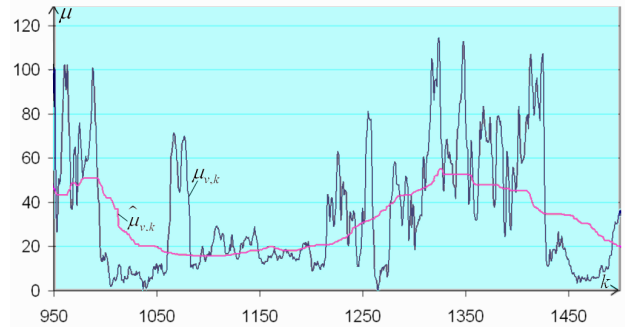


Рис. 3. АКФ столбцов и эталонное значение

Найдем мультипликативный коэффициент коррекции для k -го столбца как

$$g_k = \text{med}_v \left(\sqrt{\hat{\mu}_{vk} / \mu_{vk}}, v = \overline{1, V} \right). \quad (17)$$

Здесь также (рис. 4) медианный фильтр выбирает по множеству фрагментов из $g_{kv} = \sqrt{\hat{\mu}_{vk} / \mu_{vk}}$ наиболее представительный отсчет мультипликативного коэффициента.

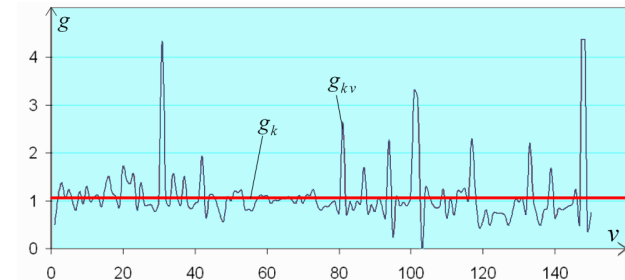


Рис. 4. Изменение g_{kv} от v .

При определении аддитивных коэффициентов для всех фрагментов используется единый мультипликативный коэффициент g_k :

$$a_{kv} = \text{med}_s \left(\frac{1}{N} \sum_{j=1}^N b_{v,k+s,j}, s = \overline{-S, S} \right) - \frac{g_k}{N} \sum_{j=1}^N b_{vkj}, \quad (18)$$

а его итоговое значение для столбца находится как

$$a_k = \text{med}_v (a_{kv}, v = \overline{1, V}), \quad k = \overline{1, K}. \quad (19)$$

В данном случае выбор медианного фильтра определяется теми же соображениями.

Следует отметить, что рассмотренные выше этапы радиометрического выравнивания сканов и отдельных фотоприемников выполняются последовательно, но коррекция собственно изображения осуществляется только один раз по формуле

$$b_{ki}^* = g_k (r_i b_{ki} + c_i) + a_k, \quad k = \overline{1, K}, \quad i = \overline{1, I}, \quad (20)$$

где b_{ki} и b_{ki}^* - исходная и скорректированная яркости элементов k -го столбца i -го скана; r_i , c_i - коэффициенты яркостного выравнивания i -го скана, а g_k , a_k - коэффициенты коррекции структурных искажений видеоданных k -го столбца. Результат такой обработки представлен на рис. 5.

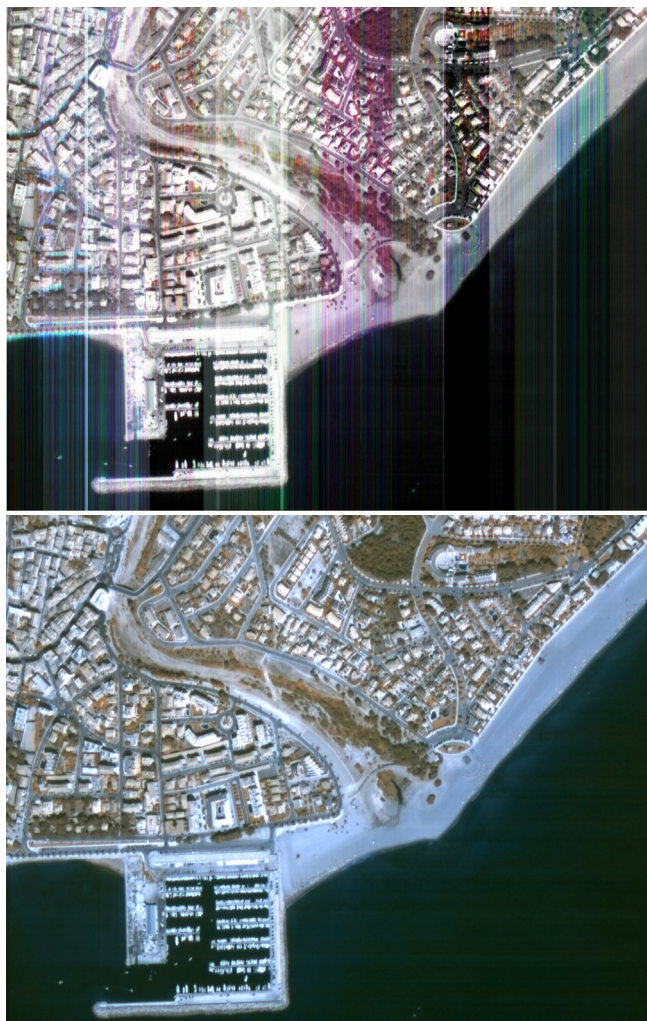


Рис. 5. Изображение до (вверху) и после коррекции

Заключение

Рассмотренные выше статистические алгоритмы яркостной коррекции изображений апробированы на реальной видеoinформации, полученной от космических систем «Ресурс-О1», «Монитор-Э» и «Ресурс-ДК». По результатам экспериментальных исследований можно сделать следующие выводы.

1. Предложенные подходы по статистической коррекции структурных искажений на изображениях, содержащих существенно отличающиеся по яркости объекты, позволяют качественно решить эту задачу, в то время как традиционно используемые алгоритмы [1, 2] оказываются неработоспособными.

2. Относительная погрешность яркостного выравнивания сканов рассмотренным выше алгоритмом не превышает 0,3%.

3. Относительная погрешность коррекции структурных искажений отдельных фотоприемников с использованием алгоритмов, основанных на нормировании функций распределения, кластеризации изображений и фильтрации неоднородности сюжета, составила порядка (0,2-0,3)%, так что остаточные искажения становятся визуально незаметными.

Рассмотренные в настоящей работе алгоритмы реализованы программно и используются в тех или иных модификациях в Научном центре оперативного мониторинга Земли и Научно-исследовательском центре космической гидрометеорологии «Планета» для обработки спутниковых изображений, получаемых от отечественных и зарубежных систем дистанционного зондирования Земли.

Литература

1. Злобин В.К., Еремеев В.В. Обработка аэрокосмических изображений. – М.: Физматлит, 2006. 286 с.
2. Еремеев В.В., Злобин В.К. Статистические алгоритмы радиометрической коррекции видеoinформации от многоэлементных сканирующих систем // Автометрия. 1995. №2. С. 78-83.
3. Зенин В.А., Кузнецов А.Е., Побаруев В.И. Алгоритм радиометрической коррекции изображений с неоднородным сюжетом, полученных от космического аппарата «Ресурс-ДК» // Вестник РГПУ. Вып. 23, Рязань 2008. С. 43-48.

RADIOMETRIC ALIGNMENT OF VIDEO DATA FROM MULTIMATRIX SCANNERS

V.V. Eremeev, V.A. Zenin

The decision of a problem of statistical radiometric correction of structural distortions on the satellite images of a terrestrial surface formed by means of multimatrix video gauges is presented. Algorithms of correction of the images containing essentially non-uniform objects on brightness when known algorithms appear not efficient are considered. Approaches on alignment of video data from various photoreception matrixes are offered and elimination of the structural noise caused by distinction of sensitivity of separate photoreception elements. Results of processing of real pictures are resulted.

ФОРМИРОВАНИЕ ЦВЕТНЫХ СНИМКОВ ПО СПУТНИКОВЫМ ИЗОБРАЖЕНИЯМ СРЕДНЕГО И ВЫСОКОГО ПРОСТРАНСТВЕННОГО РАЗРЕШЕНИЯ

Кузнецов А.Е., д.т.н, зам.директора НИИ «Фотон» РГРТУ,
Светелкин П.Н., м.н.с. НИИ «Фотон» РГРТУ, foton@rgta.ryazan.ru

Ключевые слова: дистанционное зондирование, спутниковый снимок, сканер, пространственное разрешение, цветное изображение, модель комплексирования, триангуляция.

Введение

В последнее время в системах дистанционного зондирования Земли среднего (КА «Метеор-М») и высокого пространственного разрешения (КА «Ресурс-ДК») проектирование съемочных устройств (СУ) осуществляется по принципу пространственно распределенной регистрации поступающего электромагнитного излучения. В соответствии с этим принципом в фокальной плоскости сканера на некотором расстоянии друг от друга устанавливаются три оптико-электронных преобразователя (ОЭП), каждый из которых обеспечивает наблюдение подстилающей поверхности в определенном спектральном диапазоне. В результате съемки формируются три разновременных спектрозональных снимка $B_k = \{b_k(m_k, n_k)\}$, $m_k = \overline{0, M_k - 1}$, $n_k = \overline{0, N_k - 1}$, $k = \overline{1, 3}$, пространственно смещенные друг относительно друга.

Подобный принцип съемки с одной стороны позволяет, во-первых, упростить и удешевить конструкцию сканера и, во-вторых, получать изображения B_k с более высокими радиометрическими характеристиками. Но с другой стороны возникают дополнительные требования к наземной обработке получаемой видеoinформации, связанные прежде всего с формированием цветосинтезированных снимков. В настоящей работе рассматриваются алгоритмы получения цветных снимков, связанные с геометрическим и радиометрическим комплексированием исходных изображений в системе координат базового снимка (снимка, сформированного вторым по счету ОЭП).

Модели геометрического комплексирования изображений

Один из подходов геометрического совмещения спектрозональных снимков базируется на аналитической модели координатного соответствия, параметры которой можно рассчитать на основе уравнений географической привязки изображений B_k . Эти уравнения составляются на основе параметров орбитального положения КА во время съемки и его ориентации при известных параметрах внутреннего ориентирования съемочного устройства [1] и представляется в виде:

$$\varphi_k = F_k(m_k, n_k), \lambda_k = \Phi_k(m_k, n_k), k = \overline{1, 3}, \quad (1)$$

Рассматриваются модели геометрического и радиометрического комплексирования спутниковых снимков среднего и высокого разрешения с целью получения цветных изображений.

где (φ_k, λ_k) – геодезические координаты пикселя $b_k(m_k, n_k)$. Тогда аналитическая модель геометрического соответствия между базовым и совмещаемым изображением будет выглядеть следующим образом:

$$m_k = F_k^*(m, n), n_k = \Phi_k^*(m, n), F_k^*(m, n) = F_k^{-1}[F(m, n), \Phi(m, n)], \\ \Phi_k^*(m, n) = \Phi_k^{-1}[F(m, n), \Phi(m, n)], m = m_2, n = n_2, k = \overline{1, 3}, \quad (2)$$

где F_k^{-1} , Φ_k^{-1} – обратные функции географической привязки элементов изображения.

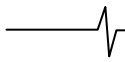
Основная сложность построения аналитической модели координатного соответствия заключается в том, что на основе модели орбитальной съемки обратные функции F_k^{-1} , Φ_k^{-1} получить не удастся. Поэтому для определения зависимости планарных (m_k, n_k) от геодезических координат (φ_k, λ_k) снимка воспользуемся следующим итерационным алгоритмом. Сначала определим начальное приближение планарных координат (m_{k0}, n_{k0}) искомой точки, в качестве которого могут, например, выступать координаты центра кадра: $m_{k0} = M_k/2$, $n_{k0} = N_k/2$. Затем на основе формулы приращения функции нескольких переменных применительно к прямым зависимостям запишем:

$$\varphi_k = \varphi + \frac{\partial F_k}{\partial m_k} \Delta m_k + \frac{\partial F_k}{\partial n_k} \Delta n_k, \\ \lambda_k = \lambda + \frac{\partial \Phi_k}{\partial m_k} \Delta m_k + \frac{\partial \Phi_k}{\partial n_k} \Delta n_k. \quad (3)$$

Выражая Δm_k , Δn_k через φ_k , λ_k , получим поправки к планарным координатам искомой точки для i -го шага итерации:

$$\Delta m_{ki} = \frac{\left((\lambda_{ki} - \lambda) \frac{\partial F_k}{\partial n_k} - (\varphi_{ki} - \varphi) \frac{\partial \Phi_k}{\partial n_k} \right)}{\left(\frac{\partial F_k}{\partial m_k} \frac{\partial \Phi_k}{\partial n_k} - \frac{\partial F_k}{\partial n_k} \frac{\partial \Phi_k}{\partial m_k} \right)}, \\ \Delta n_{ki} = \frac{\left((\varphi_{ki} - \varphi) \frac{\partial \Phi_k}{\partial m_k} - (\lambda_{ki} - \lambda) \frac{\partial F_k}{\partial m_k} \right)}{\left(\frac{\partial F_k}{\partial m_k} \frac{\partial \Phi_k}{\partial n_k} - \frac{\partial F_k}{\partial n_k} \frac{\partial \Phi_k}{\partial m_k} \right)}. \quad (4)$$

Частные производные в формулах (4) вычислим через конечные разности, тогда географические координата-



ты для i -ой итерации определим как $\varphi_{ki} = F_k(m_{ki}, n_{ki})$, $\lambda_{ki} = \Phi_k(m_{ki}, n_{ki})$. На каждой итерации выполним уточнение географических координат искомой точки $m_{k,i+1} = m_{ki} + \Delta m_{ki}$, $n_{k,i+1} = n_{ki} + \Delta n_{ki}$.

Условием выхода из итерационного процесса служит достижение рассогласования в плоскостных координатах незначимых величин $\Delta m_{ki} \leq \delta_m$, $\Delta n_{ki} \leq \delta_n$, где δ_m , δ_n выбираются в зависимости от необходимой точности решения поставленной задачи и обычно составляют десятые или сотые доли пикселя. С учетом описанного выше итерационного алгоритма схема вычисления Z значений функций $F_k^*(m, n)$, $\Phi_k^*(m, n)$ будет выглядеть так:

$$(m_z, n_z) \xrightarrow{F_z, \Phi_z} (\varphi_z, \lambda_z) \xrightarrow{F_k^{-1}, \Phi_k^{-1}} (m_{kz}, n_{kz}), \quad z = \overline{1, Z}. \quad (5)$$

Таким образом, получив алгоритм вычисления значений зависимостей (2), можно выполнить совмещение снимков и получить новые спектрональные изображения $D_k = \{d_k(m, n)\}$ в единой системе координат. Однако, пересчет каждого пикселя исходного снимка в систему координат базового займет значительное время. Поэтому для ускорения процесса совмещения снимков по аналитической модели (2) функции F_k^* , Φ_k^* представляются в кусочно-билинейном виде. В таком виде аналитическая модель применяется для комплексирования снимков от МСУ-50/100, устанавливаемом на планируемом к запуску КА «Метеор-М» №1.

Для геометрического совмещения снимков высокого разрешения модели (2) недостаточно, т.к. из-за ошибок определения параметров внешнего и внутреннего ориентирования и рельефа местности погрешности вычисления координат одноименных точек снимков могут достигать десятков и сотен пикселей. Поэтому при совмещении снимков высокого пространственного разрешения используется двухэтапный подход, в соответствии с которым вначале выполняется оперативное совмещение снимков по полиномиальной модели с учетом углового движения КА, а затем прецизионное совмещение с использованием триангуляционной модели, учитывающей рельефные искажения.

Параметры полиномиальной модели определяются на основе координат одноименных точек спектрональных снимков. Для их поиска в условиях неизвестного пространственного рассогласования изображений B_k предлагается пирамидальный алгоритм поиска одноименных сюжетов. В соответствии с этим алгоритмом в углах снимка определяются информативные фрагменты, подходящие для дальнейшего корреляционно-экстремального отождествления (КЭО). Далее, формируется пирамида разномасштабных фрагментов $\tilde{B}_{k\phi S}^*$, $s = \overline{1, S}$,

$$B_{k\phi}^* = \tilde{B}_{k\phi 1}^* \xrightarrow{\Psi_2} \tilde{B}_{k\phi 2}^* \xrightarrow{\Psi_3} \dots \xrightarrow{\Psi_S} \tilde{B}_{k\phi S}^*, \quad (6)$$

где Ψ – оператор масштабного преобразования с коэффициентом μ_s . Идея алгоритма состоит в том, что если на нижнем уровне пирамиды наблюдается рассогласование ℓ пикселей, то на верхнем уровне пирамиды это рассогласование равноется

$$\ell / \prod_{s=2}^S \mu_s \text{ пикселей, т.е. можно значительно}$$

сократить время вычисления корреляционной функции за счет сокращения области поиска на следующем уров-

не пирамиды. Далее, определив на основе КЭО взаимные смещения на нулевом уровне пирамиды, строится полиномиальная модель взаимного соответствия спектрональных изображений:

$$\begin{aligned} \widehat{B}_k &= \{ \widehat{b}_k(m, n) = b_k(m_k, n_k) \}, \\ m_k &= P_{mk}(m, n) = p_{0k} + p_{1k}m + p_{2k}n, \\ n_k &= P_{nk}(m, n) = p_{3k} + p_{4k}m + p_{5k}n, \end{aligned} \quad (7)$$

где \widehat{B}_k – геометрически преобразованные в систему координат (m, n) снимки B_k . При этом коэффициенты полиномов P_{mk} , P_{nk} определяются по идентифицированным координатам одноименных точек с использованием метода наименьших квадратов.

По модели (7) можно совместить снимки высокого разрешения равнинной местности с точностью 2-3 пикселя. В процессе исследований выяснилось, что на взаимные рассогласования изображений высокого разрешения также влияют изменения углов ориентации КА во время съемки, и модель (7) не всегда оказывается адекватной. На протяженных маршрутах, при наличии значительного углового движения, взаимные рассогласования снимков могут достигать десятков пикселей. В связи с этим модель (7) с учетом изменения углов ориентации КА представляется в виде:

$$\begin{aligned} m_k &= p_{k0} + p_{k1}[m + C_{mk}(n)] + p_{k2}[n + C_{nk}(n)], \\ n_k &= p_{k3} + p_{k4}[m + C_{mk}(n)] + p_{k5}[n + C_{nk}(n)], \quad k = 1, 3, \end{aligned} \quad (8)$$

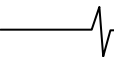
где $C_{mk}(n)$, $C_{nk}(n)$ – взаимные смещения совмещаемых снимков B_k относительно базового в строчном и кадровом направлениях, вызванные угловым движением КА. Эти смещения определяются на основе угловых скоростей КА:

$$\begin{aligned} C_{mk}(n) &\approx \frac{f[\Delta\omega(t, \Delta t_k) - \Delta\omega(t_0, \Delta t_k)]}{r}, \\ C_{nk}(n) &\approx \frac{f[\Delta\alpha(t, \Delta t_k) - \Delta\alpha(t_0, \Delta t_k)]}{r}, \end{aligned} \quad (9)$$

где f – фокусное расстояние съемочного устройства, r – размер фотоэлемента линейки ПЗС, $\Delta\omega(t, \Delta t_k)$, $\Delta\alpha(t, \Delta t_k)$ – изменение угла крена и тангажа от момента времени t_0 до Δt_k , t_0 – момент времени формирования первой строки изображения базового ОЭП, Δt_k – временной интервал между съемкой одноименных объектов местности соседними ОЭП.

На рис.1 показан пример совмещения маршрута съемки «Париж» по модели (7) (рис. 1, а) и по модели (8) (рис. 1, б). В центре снимка (рис. 1, а) видны рассогласования достигающие сотни пикселей, на рис. 1, б эти рассогласования устранены.

После совмещения по модели (8) могут остаться взаимные рассогласования, обусловленные главным образом рельефом местности. Для устранения указанных рассогласований используется триангуляционная модель геометрического соответствия, в соответствии с которой плоскость совмещаемого снимка разбивается на треугольные фрагменты, в пределах которых координатная обработка выполняется по линейному закону. Например, для изображения B_1 в пределах треугольника с порядковым номером u соотношения выглядят следующим образом:



$$m_1 = a_{u_0} + a_{u_1}m + a_{u_2}n, \quad n_1 = a_{u_3} + a_{u_4}m + a_{u_5}n. \quad (10)$$

Вершинами триангуляционной решетки являются идентифицированные на изображениях $\hat{B}_1, \hat{B}_2, \hat{B}_3$ одноименные точки. Изначально точки на базовом снимке формируются в узлах регулярной прямоугольной решетки с шагом, позволяющим аппроксимировать с заданной точностью соотношения (8). При этом их образы на изображениях \hat{B}_1 и \hat{B}_3 определяются с использованием аналитико-регрессионной модели, а потом уточняются с использованием аппарата КЭО. Разбиение плоскости на треугольники

выполняется по критерию Делоне [2].

Далее триангуляционная модель автоматически загущается до требуемой точности совмещения, которая как правило составляет 0,5 пикселя. На рис. 2 показан маршрут «Тегеран», совмещенный с использованием регулярной триангуляционной сети (рис. 2, а) и сети, загущенной с учетом рельефа местности (рис. 2, б). Из рис. 2, б видно, что триангуляционная сеть загущается на горных районах маршрута, а на относительно ровных участках остается без изменения.

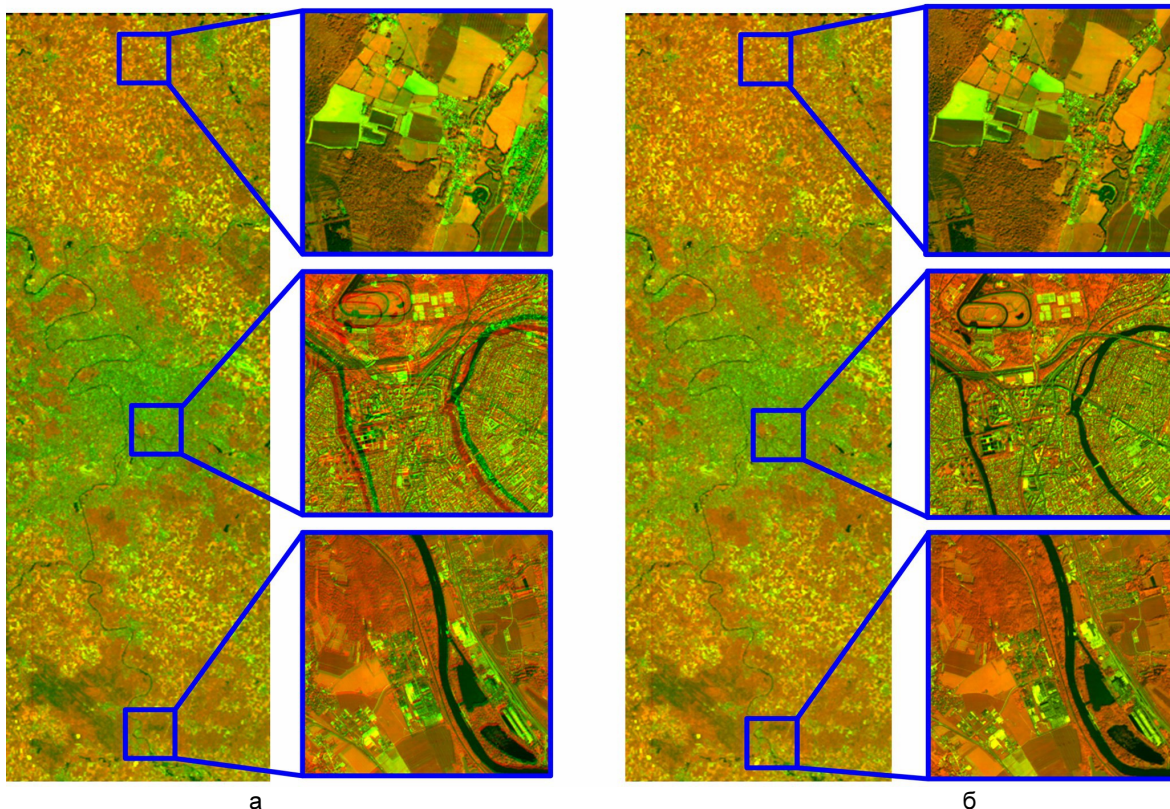


Рис. 1. Пример совмещения двух снимков: а – без учета изменения углов ориентации КА, б – с учетом изменения углов ориентации

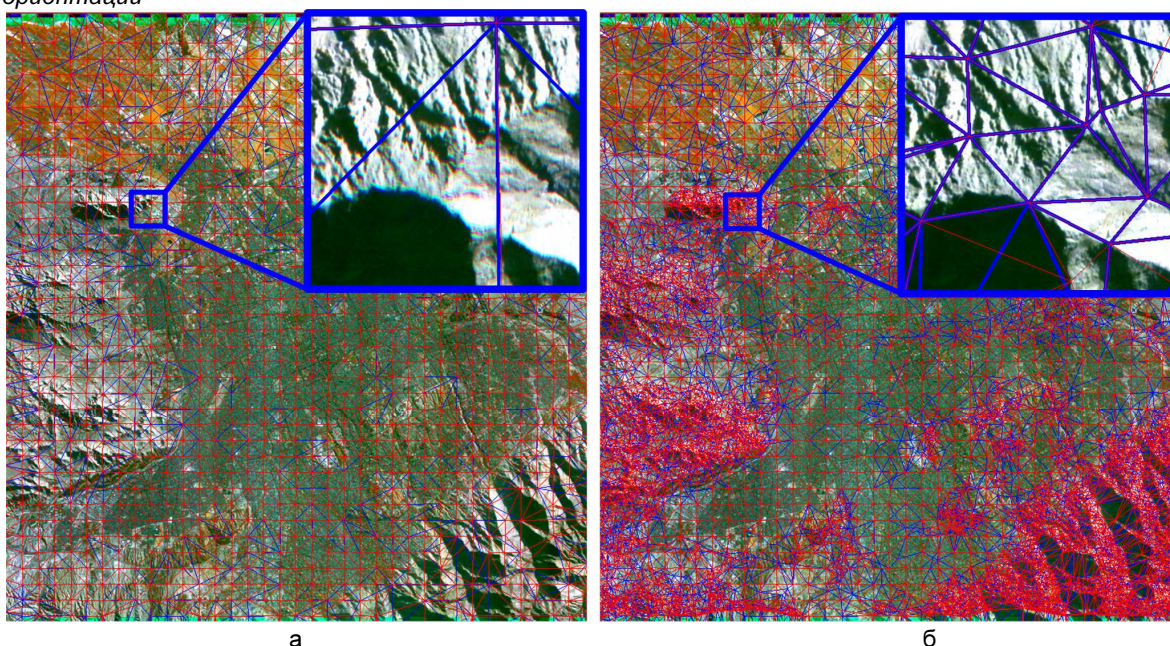


Рис. 2. Маршрут «Тегеран», совмещенный по триангуляционной модели: а – регулярная сетка, б – сетка, загущенная с учетом рельефа местности

Для повышения надежности корреляционно-экстремальной идентификации, в частности, для отбраковки точек, попавших на движущиеся объекты, предложен следующий алгоритм.

Установлено, что перепады высот снимаемых объектов влияют на рассогласования в кадровом направлении и практически не оказывают влияния на рассогласования в строчном направлении. Соответственно строчные смещения совмещаемого снимка относительно базового в одноименных точках местности распределены по нормальному закону, а одноименные точки, попавшие на движущиеся объекты, можно отбраковать по их строчным смещениям по критерию «3σ». На рис. 3 приведен пример смещений Δm в строчном направлении. На рис. 3,а видно два пика, которые образованы точками, попавшими на движущиеся объекты. На рис. 3,б показан результат отбраковки одноименных точек, попавших на движущиеся объекты.

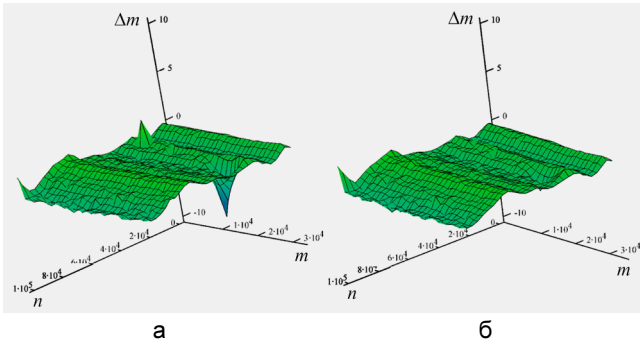


Рис. 3. Графики строчных смещений одноименных точек: а – смещения до отбраковки движущихся объектов, б – после отбраковки

Радиометрическое комплексирование спектрозональных снимков

Съемка тремя ОЭП не всегда выполняется в спектральных диапазонах, соответствующих «красной», «зеленой» и «синей» области спектра. Например, на КА «Ресурс-ДК» вместо «зеленой» может выполняться съемка в широкой панхроматической области спектра, поэтому при создании цветосинтезированного изображения необходимо восстановить недостающие компоненты.

Если в режиме RGB-съемки компоненты цветного снимка остаются неизменными, то в режиме RPB-съемки компоненты будут формироваться в соответствии с выражениями:

$$D_R = \hat{B}_3, D_G = 3B_2 - \hat{B}_1 - \hat{B}_3, D_B = \hat{B}_1 \quad (11)$$

где D_R, D_G, D_B – красная, зеленая и синяя компоненты цветосинтезированного изображения D , соответственно, в режиме PGB-съемки аналогично:

$$D_R = 3\hat{B}_3 - \hat{B}_1 - B_2, D_G = B_2, D_B = \hat{B}_1. \quad (12)$$

При отсутствии снимка от одного из ОЭП и наличии снимка панхроматического диапазона компоненты цветного снимка вычисляются на основе предположения, что гистограмма естественной сцены аппроксимируется полиномом третьей степени.

BP-съемка:

$$D_R = \beta_0 + \beta_1 \hat{B}_1 + \beta_2 \hat{B}_1^2 + \beta_3 \hat{B}_1^3, D_G = 3B_2 - \hat{B}_1 - D_R, D_B = \hat{B}_1, \quad (13)$$

где $\beta_0, \beta_1, \beta_2, \beta_3$, – коэффициенты полинома 3-й степени, вычисленные по гистограмме изображения \hat{B}_1 .

PR-съемка:

$$D_R = \hat{B}_3, D_G = 3B_2 - D_B - \hat{B}_3, D_B = \beta_4 + \beta_5 \hat{B}_3 + \beta_6 \hat{B}_3^2 + \beta_7 \hat{B}_3^3, \quad (14)$$

где $\beta_4, \beta_5, \beta_6, \beta_7$, – коэффициенты полинома 3-й степени, вычисленные по гистограмме изображения B_3 .

GP-съемка:

$$D_R = 3\hat{B}_3 - D_B - B_2, D_G = B_2, D_B = \beta_8 + \beta_9 B_2 + \beta_{10} B_2^2 + \beta_{11} B_2^3, \quad (15)$$

где $\beta_8, \beta_9, \beta_{10}, \beta_{11}$ – коэффициенты полинома 3-й степени, вычисленные по гистограмме изображения B_2 .

Вследствие того, что спектрозональные снимки являются разновременными, движущиеся объекты представляются на них как три разноцветных объекта (рис.4, а). Для устранения влияния движущихся объектов на качество снимка был предложен модуляционный алгоритм радиометрического комплексирования, который функционирует при наличии панхроматического снимка. Работу модуляционного алгоритма можно разделить на два этапа. На первом этапе выполняется «размытие» спектрозональных компонент в соответствии с формулами

$$\hat{B}_1^* = \hat{B}_1 \otimes h_{j \times j}, \hat{B}_3^* = \hat{B}_3 \otimes h_{j \times j}, \quad (16)$$

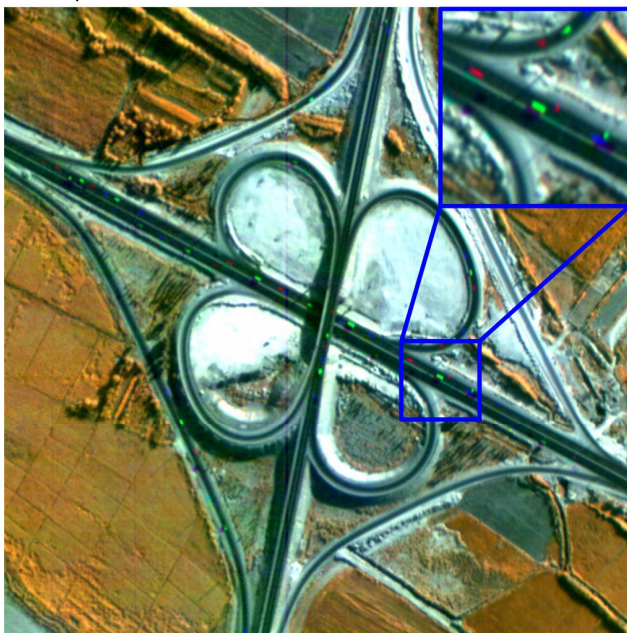
где \otimes – операция свертки с маской размытия h размером $j \times j$ элементов. На втором этапе получаем компоненты цветосинтезированного изображения с использованием модуляции панхроматической компоненты синей и красной компонентами по формулам:

$$D_R = \frac{3\hat{B}_3^*}{B_2 + \hat{B}_1^* + \hat{B}_3^*} B_2, D_G = B_2, D_B = \frac{3\hat{B}_1^*}{B_2 + \hat{B}_1^* + \hat{B}_3^*} B_2. \quad (17)$$

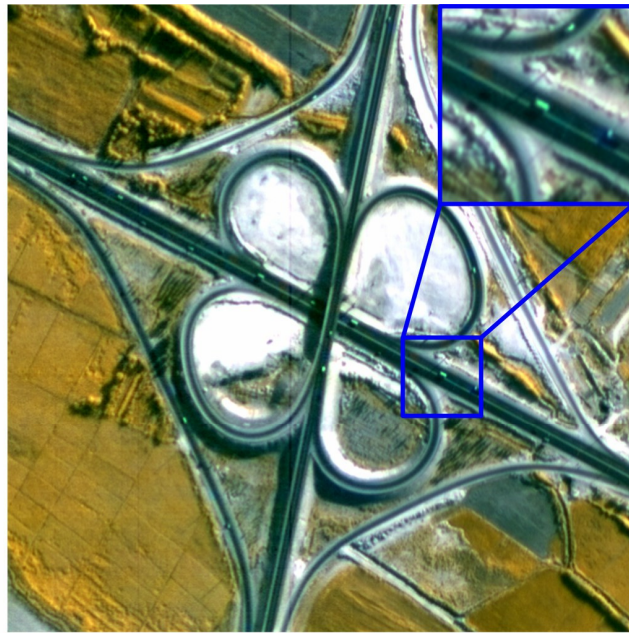
Указанный алгоритм позволяет также устранить остаточные рассогласования 1-2 пикселя от высотных объектов (башен, небоскребов). Результат работы модуляционного алгоритма представлен на рис. 4,б.

Заключение

Рассмотренные в статье алгоритмы геометрического и радиометрического комплексирования легли в основу программных модулей создания цветосинтезированных изображений в ряде программных комплексов обработки видеоданных от отечественных КА, таких как «Метеор-М» №1 и «Ресурс-ДК». При этом для снимков среднего разрешения используется аналитическая модель геометрического соответствия, а в случае совмещения снимков высокого разрешения применяется аналитико-регрессионная и триангуляционная модели. Эффективность предложенных решений подтверждается многолетней успешной эксплуатацией разработанных программных комплексов.



а



б

Рис. 4. Пример маскирования движущихся объектов: а – исходный снимок, б – снимок, полученные с использованием модуляционного алгоритма

Литература

1. Обработка изображений в геоинформационных системах: Учебное пособие / В.К.Злобин, В.В.Еремеев, А.Е.Кузнецов. Рязан. Гос. Радиотехн. Университет, Рязань, 2006. 264 с.
2. Скворцов А.В. Триангуляция Делоне и ее применение. – Томск: Изд-во Том. ун-та, 2002. – 128 с.
3. Бакланов А.И. Системы наблюдения и мониторинга: учебное пособие / А.И. Бакланов. - М.: БИНОМ. Лаборатория знаний, 2009. 234с. : ил.

FORMING COLOR IMAGES FROM REMOTE SENSING DATA MEDIUM AND HIGH SPATIAL RESOLUTION

Kuznetsov A.E., Svetelkin P.N.

Models geometric and radiometric fusion of remote sensing data medium and high spatial resolution for getting color images is consider

Работа проводилась при финансовой поддержке Федерального агентства по науке и инновациям (госконтракт № 02.740.11.0002).

ИНФОРМАЦИОННОЕ ПИСЬМО

Уважаемый коллега!

Приглашаем Вас принять участие в работе 12-ой Международной конференции и выставки «Цифровая обработка сигналов и ее применение – DSPA'2010», которая состоится в Москве в ИПУ РАН 24-26 марта 2010 года

ОРГАНИЗАТОРЫ:

- Российское научно-техническое общество радиотехники, электроники и связи им. А.С. Попова
- IEEE Signal Processing Society
- Российская секция IEEE
- Институт радиотехники и электроники РАН
- Институт проблем управления РАН
- Институт проблем передачи информации РАН
- Московский научно-исследовательский телевизионный институт
- Компания AUTEX Ltd. (АВТЭКС)

ОСНОВНЫЕ НАПРАВЛЕНИЯ РАБОТЫ КОНФЕРЕНЦИИ:

- Теория сигналов и систем
- Теория и методы ЦОС
- Цифровая обработка многомерных сигналов
- Цифровая обработка речевых и звуковых сигналов
- Цифровая обработка изображений
- ЦОС в системах телекоммуникаций
- ЦОС в радиотехнических системах
- ЦОС в системах управления и робототехники
- Цифровая обработка измерительной информации
- Нейрокомпьютерная обработка сигналов и изображений
- Цифровое телевидение и телевидение
- Цифровое радиовещание
- ЦОС в системах защиты информации
- Проектирование и техническая реализация систем ЦОС
- ЦОС в открытых системах
- Проблемы подготовки специалистов в области ЦОС

УДК 004.932.2: 528.88

ОЦЕНКА ВЫСОТЫ ОБЛАЧНОСТИ ПО ДАННЫМ КОМПЛЕКСА МНОГОЗОНАЛЬНОЙ СКАНЕРНОЙ СЪЕМКИ КОСМИЧЕСКОГО АППАРАТА «МЕТЕОР-М»

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

Пошехонов В.И., аспирант НИИ «Фотон» РГРТУ, foton@rgta.ryazan.ru

Ключевые слова: космический аппарат, сканерная съемка, метеообстановка, спутниковые данные, стереоизображение, ПЗС-линейка, цифровая карта, стереообработка, вычислительный процесс.

Введение

Одной из важных прикладных гидрометеорологических задач является оперативное определение метеообстановки в заданном регионе по спутниковым данным, включая оценку высоты облачных образований. Для этого в зарубежных космических системах применяется ряд подходов. Первый подход основан на измерении яркости отраженного излучения в определенных диапазонах инфракрасной области спектра, что позволяет косвенно определить высоту облачности с использованием дополнительных данных о коэффициенте ее излучения и вертикальном градиенте температуры [1]. Второй подход предполагает использование активных сенсоров, таких как, например, радар миллиметрового диапазона CloudSat или лидар CALIOP, позволяющих получать вертикальные профили облачности в узкой полосе наблюдения спутника [2].

И, наконец, третий подход расчета высоты облачности, основанный на использовании стереоскопических наблюдений, применяется как для изображений от геостационарных, так и других спутников [3]. Принцип обработки стереоизображений, основанный на базовых геометрических зависимостях и определении на снимках координат одноименных объектов, является единственным методом расчета высоты облачности по данным от пассивных сенсоров, не зависящим от каких-либо дополнительных условий.

В этой связи весьма актуальными являются исследования, связанные с изучением возможности использования информации от запланированного к запуску Российского гидрометеорологического космического аппарата (КА) «Метеор-М» [4] для определения параметров облачности. Одним из приборов дистанционного зондирования, установленных на данном спутнике, является комплекс многозональной сканерной съемки (КМСС). Несмотря на то, что съемочные устройства, входящие в состав КМСС, не разрабатывались для стереоизмерений [5], их конструктивные особенности приводят к тому, что формирование изображений одних и тех же объектов в разных спектральных диапазонах выполняется с временной задержкой и под разными от надира углами. Отсюда возникают вопросы, связанные с

Исследуется возможность оценки высоты облачности по изображениям от комплекса многозональной сканерной съемки, установленного на космическом аппарате «Метеор-М» №1. Рассматриваются технология и алгоритмы стереообработки видеоданных на современной вычислительной технике.

исследованием точностных характеристик получаемых оценок высоты облачности стереофотограмметрическими методами и разработкой технологии обработки спектральнозональных изображений с использованием современной вычислительной техники с целью оперативного формирования выходных информационных продуктов.

Исследование возможности стереообработки данных КМСС

Аппаратура КМСС включает три многозональных съемочных устройства (МСУ): одно МСУ-50 и два МСУ-100, идентичные по принципу функционирования и конструктивным решениям, отличающиеся только фокусным расстоянием объективов. Каждый сканер формирует три спектральнозональных изображения с помощью трех ПЗС-линеек, установленных в фокальной плоскости с небольшим смещением вдоль направления полета КА, что приводит к отклонению от надира плоскости визирования и появлению на изображениях стереоэффекта. Рассмотрим возможность измерения высоты облачности по формируемому первой и третьей ПЗС-линейками изображениям, имеющим наибольшую стереобазу. Для этого представим упрощенную модель съемки (рис. 1), в которой игнорируются вращение Земли и отклонение визирных лучей поперек трассы полета КА, орбита предполагается круговой, а Земля – сферической радиусом R .

Облачность в точке P сканируется первой ПЗС-линейкой из точки M и проецируется на поверхность Земли в точку P' . Через интервал времени Δt происходит сканирование облачности другой ПЗС-линейкой под углом от надира ν и проецируется в точку P'' . Угол между визирным лучом и надиром для первого и третьего каналов камеры МСУ-50 определим из соотношения $\nu_{M50} = \arctg(a/f_{M50})$, где $a = 15.24$ мм – расстояние между соседними ПЗС-линейками в фокальной плоскости, $f_{M50} = 50$ мм – фокусное расстояние объектива МСУ-50. Поскольку камеры МСУ-100 устанавливаются под углом $\gamma = 14^\circ$ относительно местной вертикали для увеличения суммарной полосы захвата, сканирование подспутниковой точки выполняется крайними элементами ПЗС-линейки, поэтому

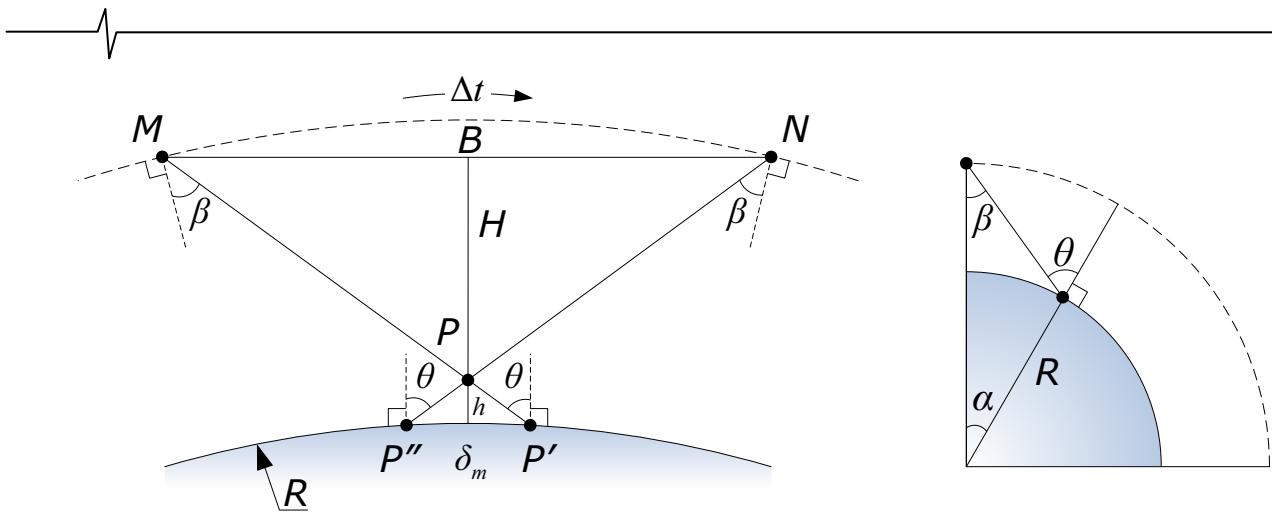


Рис. 1. Геометрия съемки вдоль направления полета спутника

$$\beta_{M100} = \arctg\left(\frac{a}{f_{M100}/\cos\gamma}\right),$$

где $f_{M100} = 100$ мм – фокусное расстояние объектива МСУ-100.

Тогда с учетом кривизны Земли и высоты полета спутника, равной 832 км, отклонение визирных лучей от зенита на поверхности Земли, которое определяется исходя из соотношения $\theta = \beta + \alpha$ (рис. 1), составит $\pm 19.25^\circ$ для МСУ-50 и $\pm 9.52^\circ$ для МСУ-100.

Так как высота облачности $h \ll R$, то линию, соединяющую точки P' и P'' на поверхности Земли, можно считать прямой. Тогда диспаратность δ_m , равная расстоянию $P'P''$, пропорциональна высоте h ,

$$h = \frac{\delta_m}{2 \operatorname{tg}\theta} = \frac{\delta_m}{(B/H)},$$

где (B/H) – отношение базы стереосъемки к высоте.

Оценим точность определения диспаратности σ_δ по данным КМСС. Погрешность измерения координат одноименных точек σ_u примем равной разрешению снимков, которое составляет 120 м для МСУ-50 и 60 м для МСУ-100. Ожидаемая точность геопривязки изображений составляет порядка 1 пикселя МСУ-100 [5], что приведет к дополнительной погрешности $\sigma_z = 60$ м для обоих устройств. Таким образом, точность

$$\sigma_\delta = \sqrt{\sigma_u^2 + \sigma_z^2} \text{ равна } 134.16 \text{ м для МСУ-50 и } 84.85 \text{ м для МСУ-100.}$$

Погрешность определения высоты σ_h , равная

$$\sigma_h = \frac{\sigma_\delta}{(B/H)},$$

зависит также от отношения базы стереосъемки к высоте, которое максимально для пары изображений первого и третьего каналов и составляет $(B/H)_{M50} = 0.698$, $(B/H)_{M100} = 0.335$. В результате получим $\sigma_h = 192.13$ м для МСУ-50 и $\sigma_h = 252.99$ м для МСУ-100.

Заметим, что съемка одних и тех же объектов в первом и третьем каналах выполняется с временной задержкой Δt , равной 76.29 с для МСУ-50 и 36.83 с для МСУ-100. Предположим, что облачность движется параллельно поверхности Земли и ее высота не изменя-

ется со временем. Тогда в случае, если скорость движения составляет v_m вдоль трассы полета КА, это приведет к ошибке измерения диспаратности $\Delta\delta_m = v_m \Delta t$ и дополнительной погрешности определения высоты $\sigma_h(v)$, которая будет линейно зависеть от v_m и составит порядка 110 м на каждый 1 м/с скорости для обоих устройств. При этом данная ошибка может быть устранена в том случае, если значение v_m известно из какого-либо другого источника, например, по данным мониторинга с геостационарного спутника [6]. С другой стороны, известно, что скорость движения облачности вдоль меридианов, которая примерно совпадает с трассой движения КА «Метеор-М» с наклоном орбиты $i = 98^\circ$, обычно гораздо меньше, чем в направлении параллелей, и составляет в большинстве случаев на много меньше 15 м/с [7], поэтому при оценке погрешности $\sigma_h(v)$ используем значение $v_m = 3$ м/с.

Отметим, что скорость облачности v_n в направлении, перпендикулярном движению КА, можно определить по смещению одноименных точек δ_n в соответствующем направлении, $v_n = \delta_n / \Delta t$. Отсюда, используя полученную ранее оценку погрешности измерения диспаратности, точность определения скорости σ_{v_n} составит 1.76 м/с для МСУ-50 и 2.30 м/с для МСУ-100.

Сравним точность, достигаемую зарубежными спутниковыми системами определения высоты верхней границы облачности [8], с полученными оценками для аппаратуры КМСС (табл. 1). Точность, которую обеспечивает спектро радиометр MODIS, составляет порядка 1 км для облаков на средней и большой высоте, однако может ухудшаться до 3 км на малых высотах. Погрешность определения высоты аппаратурой стереосъемки MISR достигает 600 м по сравнению с результатами наземных измерений. Таким образом, геометрия съемки МСУ-50/100 позволяет получать более точные оценки высоты облачного покрова, что существенно расширяет область применения информации КМСС. Съёмочное устройство МСУ-50 при большей полосе захвата, составляющей 940 км, обеспечивает более высокую точность определения высоты облачности и больше подходит для оперативной обработки за счет меньшего объема данных по отношению к изображениям МСУ-100.

Таблица 1

Камера	Разрешение, м	B/H	Δt , с	σ_h , м	$\sigma_h(v)$, М, при $v_m = 3$ м/с	σ_{v_n} , м/с
МСУ-50 (канал 1-3)	120	0.698	76.29	192.13	327.74	1.76
МСУ-100 (канал 1-3)	60	0.335	36.83	252.99	329.46	2.30

Схема обработки данных КМСС

Схема обработки данных МСУ-50 с целью формирования цифровых карт высот облачного покрова для заданного региона предусматривает последовательное выполнение следующих шагов.

Шаг 1. Задачей первого шага обработки пары спектральных снимков $I_1(m_1, n_1)$ и $I_2(m_2, n_2)$, $m_1 = m_2 = 1, M$, $n_1 = n_2 = 1, N$, является устранение систематических взаимных искажений, обусловленных вращением и кривизной Земли, не стабильностью углов ориентации КА и др. После выполнения этого этапа смещения координат одноименных точек, вызванные высотой облачности h , будут присутствовать только в кадровом направлении, а строчные смещения обусловлены скоростью движения v_n . Совместим снимки путем геометрического трансформирования снимка I_2 в плоскость снимка I_1 , используя уравнения геопривязки изображений,

$$I_i : \varphi_i = \Phi_i(\mathbf{R}_i, \mathbf{r}_i, \mathbf{e}, m_i, n_i, f),$$

$$\lambda_i = F_i(\mathbf{R}_i, \mathbf{r}_i, \mathbf{e}, m_i, n_i, f), \quad i = 1, 2,$$

где (φ_i, λ_i) – географические координаты пикселя; \mathbf{R}_i – вектор, задающий орбитальное положение КА в момент формирования пикселя изображения $I_i(m_i, n_i)$; \mathbf{r}_i – вектор направляющих косинусов соответствующего визирного луча; \mathbf{e} – вектор параметров земного эллипсоида, f – строчная частота сканирования прибора.

Тогда уравнения взаимного соответствия снимков можно представить в виде:

$$\begin{cases} m_1 = \Phi_1^{-1}[\Phi_2(\mathbf{R}_2, \mathbf{r}_2, \mathbf{e}, m_2, n_2, f), F_2(\mathbf{R}_2, \mathbf{r}_2, \mathbf{e}, m_2, n_2, f)], \\ n_1 = F_1^{-1}[\Phi_2(\mathbf{R}_2, \mathbf{r}_2, \mathbf{e}, m_2, n_2, f), F_2(\mathbf{R}_2, \mathbf{r}_2, \mathbf{e}, m_2, n_2, f)], \end{cases}$$

где Φ_1^{-1} , F_1^{-1} – функции обратного геодезического расчета.

На основе полученных уравнений сформируем новое изображение $I_2^*(m_1, n_1)$, подлежащее дальнейшей обработке,

$$I_2(m_2, n_2) \xrightarrow[n_1 = F_1^{-1}[\Phi_2(\bullet), F_2(\bullet)]]{m_1 = \Phi_1^{-1}[\Phi_2(\bullet), F_2(\bullet)]} I_2^*(m_1, n_1).$$

Шаг 2. На втором шаге определим взаимные сдвиги одноименных объектов изображений $I_1(m, n)$ и $I_2^*(m, n)$ (здесь индексы у переменных m и n опущены), которые представим в виде карт диспарантности $D_m(m, n)$ и $D_n(m, n)$, $m = 1, M$, $n = 1, N$, задающих в пикселях смещения в кадровом и строчном направлениях соответственно. Для надежного отождествления используем алгоритм оптимизационного поиска, в соответствии с которым вначале сформируем изображение пространства диспарантности (ИПД) $C(m, n, d_m, d_n)$, где (d_m, d_n) – смещение координат точки (m, n) на изобра-

жении I_1 относительно изображения I_2^* . Многомерное ИПД задает цену сопоставления каждого пикселя базового изображения с каждым пикселем подчиненного в некоторой области поиска. Поскольку облачность на спектральных снимках представляется одинаково, то в качестве меры сходства используем нормализованную кросскорреляционную функцию K ,

$$C(m, n, d_m, d_n) = 1 - K[I_1(m, n), I_2^*(m + d_m, y + d_n)].$$

Затем для нахождения карт D_m , D_n сформулируем задачу минимизации,

$$(D_m, D_n) = \arg \min_{D_m, D_n} E(D_m, D_n),$$

$$E(D_m, D_n) = E_y(D_m, D_n) + E_z(D_m, D_n).$$

Здесь слагаемое E_y определяет, насколько карты D_m , D_n согласуются с исходными изображениями,

$$E_y(D_m, D_n) = \sum_{m, n} C(m, n, D_m(m, n), D_n(m, n)).$$

Слагаемое E_z задает требование кусочной гладкости карт D_m , D_n . Для уменьшения вычислительной сложности алгоритма функция E_z задается зависимой только от разности диспарантностей соседних пикселей,

$$E_z(D_m, D_n) = \sum_{m=1}^{M-1} \sum_{n=1}^N \rho(D_m(m, n) - D_m(m+1, n), D_n(m, n) - D_n(m+1, n)) + \sum_{m=1}^M \sum_{n=1}^{N-1} \rho(D_m(m, n) - D_m(m, n+1), D_n(m, n) - D_n(m, n+1)),$$

где ρ – функция штрафов.

Данная задача является NP-полной, поэтому нахождение глобального минимума E аппроксимируем с помощью алгоритма полуглобального сопоставления [9], в соответствии с которым выполняется 1D минимизация функции E с использованием принципа динамического программирования вдоль восьми направлений, задаваемых векторами $\mathbf{u}_k = (u_{km}, u_{kn})$, $u_{km} \in \{0, \pm 1\}$, $u_{kn} \in \{0, \pm 1\}$. Цена $L_k(m, n, d_m, d_n)$ для каждого пикселя вдоль направления k определяется рекурсивно,

$$L_k(m_j, n_j, d_m, d_n) = C(m_j, n_j, d_m, d_n) +$$

$$+ \min \{ L_k(m_{j-1}, n_{j-1}, d_m, d_n),$$

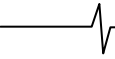
$$L_k(m_{j-1}, n_{j-1}, d_m \pm 1, d_n) + p_1,$$

$$L_k(m_{j-1}, n_{j-1}, d_m, d_n \pm 1) + p'_1, \min_{i,l} L_k(m_{j-1}, n_{j-1}, i, l) + p_2 \},$$

$$m_j = m_1 + (j-1)u_{km}, \quad n_j = n_1 + (j-1)u_{kn},$$

где $p_1 < p'_1 < p_2$ – штрафы, накладываемые при изменении диспарантности между соседними пикселями.

Полученные для всех направлений цены суммируются, после чего для каждого пикселя выбираются дис-



парантности с минимальной ценой,
 $S(m, n, d_m, d_n) = \sum_k L_k(m, n, d_m, d_n),$

$$(D_m(m, n), D_n(m, n)) = \arg \min_{d_m, d_n} S(m, n, d_m, d_n).$$

Шаг 3. На основе карт диспарантностей D_m, D_n вычислим трехмерные координаты наблюдаемых на снимках I_1 и I_2 объектов. Для этого вначале значения $(m_1 + D_m(m_1, n_1), n_1 + D_n(m_1, n_1))$ на основе функций обратного геодезического расчета преобразуем в систему координат изображения $I_2,$

$$\begin{aligned} m_1 + D_m(m_1, n_1) & \xrightarrow{m_2 = \Phi_2^{-1} \Phi_1(\cdot), F_1(\cdot)} (m_2, n_2). \\ n_1 + D_n(m_1, n_1) & \xrightarrow{n_2 = F_2^{-1} \Phi_1(\cdot), F_1(\cdot)} \end{aligned}$$

Затем для определения пространственного положения $\mathbf{P}(X, Y, Z)$ наблюдаемых объектов составим систему векторных уравнений, задающих положение визирных лучей для пикселей $I_1(m_1, n_1)$ и $I_2(m_2, n_2)$ в гринвичской геоцентрической системе координат,

$$\begin{cases} \mathbf{P} = \mathbf{r}_1 \rho_1 + \mathbf{R}_1, \\ \mathbf{P} = \mathbf{r}_2 \rho_2 + \mathbf{R}_2. \end{cases}$$

Из-за погрешностей ориентирования снимков, а также движения облачности поперек трассы полета КА, данная система уравнений в общем случае не имеет решения. Поэтому вектор искомым координат \mathbf{P} определим как среднюю точку кратчайшего отрезка $\mathbf{P}_1 \mathbf{P}_2,$ соединяющего визирные лучи,

$$\mathbf{r}_0 = \frac{\mathbf{r}_1 \times \mathbf{r}_2}{|\mathbf{r}_1 \times \mathbf{r}_2|}, \mathbf{R}_0 = \mathbf{R}_2 - \mathbf{R}_1,$$

$$\mathbf{P}_1 = \mathbf{r}_1 [(\mathbf{R}_0 \times \mathbf{r}_2) \cdot \mathbf{r}_0] + \mathbf{R}_1, \mathbf{P}_2 = \mathbf{r}_2 [(\mathbf{R}_0 \times \mathbf{r}_1) \cdot \mathbf{r}_0] + \mathbf{R}_2,$$

$$\mathbf{P} = (\mathbf{P}_1 + \mathbf{P}_2) / 2.$$

Кроме того, по длине l этого отрезка рассчитаем скорость движения v_n с целью последующей идентификации облачных образований,

$$l = |\mathbf{P}_2 - \mathbf{P}_1|, v_n = l / \Delta t, \Delta t = |m_2 - m_1| / f.$$

Далее осуществим переход от прямоугольных координат X, Y, Z к географическим φ, λ, h с помощью итерационного алгоритма по формулам:

$$\varphi = \arctg \left(\frac{Z + eL \sin \varphi}{\sqrt{X^2 + Y^2}} \right), \lambda = \arctg \frac{Y}{X},$$

$$h = X \sec \lambda \sec \varphi - L,$$

где e – эксцентриситет эллипсоида, L – радиус кривизны первого вертикала на широте φ .

В результате выполнения шага 3 для всех идентифицированных одноименных точек на изображениях I_1 и I_2 получим массивы значений высот $h(\varphi, \lambda)$ и скоростей $v_n(\varphi, \lambda)$ наблюдаемых объектов.

Шаг 4. На заключительном шаге обработки формируется итоговая карта облачности. Для этого вначале отделим облачность от объектов земной поверхности,

$$h^*(\varphi, \lambda) = \begin{cases} \Delta h(\varphi, \lambda) = h(\varphi, \lambda) - h_{\text{ЦМР}}(\varphi, \lambda), & \text{если } \Delta h(\cdot) \geq 1 \text{ км или } v_n(\varphi, \lambda) \geq 5 \text{ м/с,} \\ 0, & \text{иначе,} \end{cases}$$

где $h_{\text{ЦМР}}(\varphi, \lambda)$ – высота объектов земной поверхности, определяемая по цифровой модели рельефа (ЦМР). Для этой цели достаточно использовать такие доступные на всю территорию земной поверхности ЦМР низкого разрешения (30 угловых секунд, ≈ 1 км) как SRTM30 или GTOPO30.

Наконец, сформируем итоговую карту с цветовой легендой,

$$h^*(\varphi, \lambda) = \frac{x=F_x(\varphi, \lambda)}{y=F_y(\varphi, \lambda)} \rightarrow W : \{h^*(x, y)\} \rightarrow w(x, y),$$

где F_x, F_y – функции картографического проектирования; W – оператор цветовой кодировки значений h ; $w(x, y)$ – итоговая карта облачности.

Организация вычислительного процесса

Важным вопросом практического использования рассмотренной технологии является организация эффективных вычислений, позволяющих сформировать карту высот облачности $w(x, y)$ в максимально короткие сроки. Поскольку алгоритм построения карт диспарантностей D_m, D_n является достаточно трудоемким, то его реализация даже на современной вычислительной технике приводит к многочасовой работе. Поэтому рассмотрим возможность дополнительного сокращения временных затрат с использованием графического процессора (ГП) на примере вычисления нормализованной кросскорреляционной функции K .

Видеокарты эволюционировали от устройств с жесткой логикой до программируемых процессоров (GPU), позволяющих определять алгоритм их функционирования на отдельных стадиях конвейера обработки графических данных. Это позволило использовать их для выполнения произвольных вычислений общего назначения (general purpose GPU, GPGPU) [10]. В настоящее время ГП значительно превосходят обычные центральные процессоры (ЦП) с точки зрения вычислительной мощности. Например, ядро ГП NVIDIA GeForce GTX280 содержит 480 арифметико-логических устройств и обладает пиковой производительностью 933 GFLOP, в то время как четырехядерный ЦП Intel Core2Quad 3GHz – лишь 96 GFLOP. ГП ориентированы на максимальную пропускную способность программ с большим объемом вычислений и используют явный параллелизм по данным для увеличения производительности и скрытия задержек доступа к памяти. Рассмотренные далее алгоритмы были реализованы с использованием компилятора BrookGPU [11], который абстрагирует многие аспекты программирования графического процессора и является расширением языка C.

Базовая реализация алгоритма, которая предполагает параллельное вычисление значений корреляционной функции для каждого смещения внутри области поиска, включает выполнение нескольких шагов. На первом шаге выполняется подготовка исходных данных – вычисление математического ожидания и среднеквадратического отклонения яркостей эталона на ЦП, загрузка в графическую память исходных изображений, представленных в виде потоков значений с плавающей точкой *float*. На основном этапе работы алгоритма, выполняемом на ГП, формируется выходной поток, каждый элемент которого обрабатывается параллельно ядром (подпрограммой) вычисления корреляционной функции для заданного смещения. Таким образом, на ГП выполняются наиболее трудоемкие операции – расчет значений математиче-

ского ожидания и среднеквадратического отклонения яркостей фрагмента области поиска для заданного смещения, сумма попарных произведений значений сопоставляемых изображений. На заключительном этапе полученные коэффициенты корреляции загружаются в оперативную память компьютера для дальнейшей обработки на ЦП.

На рис. 2 синим цветом показан график ускорения корреляционного алгоритма по сравнению с реализацией на ЦП. Эталонный фрагмент составляет 32×32 пикселя, используется область поиска квадратной формы заданного размера. Корреляционный алгоритм на ЦП реализован на ассемблере с использованием команд SSE2, что дает выигрыш примерно в 2.7 раза относительно версии на языке С. Для получения временных характеристик использовалась следующая конфигурация аппаратных средств: ЦП Intel Core2Duo E6750 2.67 ГГц, видеокарта ATI Radeon X1650 XT 256 МБ, объем оперативной памяти (ОП) 3 ГБ.

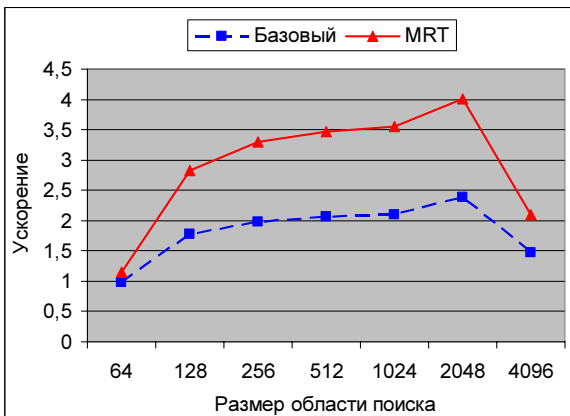


Рис. 2. Результаты ускорения корреляционного алгоритма на графическом процессоре

Низкая производительность при небольшой области поиска объясняется тем, что малый объем входных данных не позволяет эффективно загрузить все вычислительные элементы ГП. Существенное падение скорости для самой большой области поиска вызвано резким увеличением

времени считывания коэффициентов корреляции из графической памяти в основную. Это, по всей видимости, связано с тем, что при нормальном режиме функционирования видеокарты результаты обработки предназначены для вывода на экран, а не для передачи в ОП, поэтому такая операция является внештатной.

Для увеличения производительности модифицируем ядро обработки элементов выходного потока для одновременного вычисления коэффициентов корреляции для нескольких соседних пикселей вдоль строки, что позволит уменьшить требования к пропускной способности памяти, а также более эффективно организовать вычисление математического ожидания и среднеквадратического отклонения яркостей фрагментов области поиска. Для этого используем множественные цели рендеринга (multiple render targets, MRT), которые абстрагируются компилятором в виде нескольких выходных потоков. Таким образом, за счет использования ГП удается повысить быстродействие корреляционного алгоритма до 4 раз. При этом общее время обработки изображений размером 8000×8000 пикселей составляет порядка 12 минут, что вполне удовлетворяет потребностям практики.

Заключение

Выполненный анализ геометрии съемки камерами МСУ-50, МСУ-100 показал принципиальную возможность стереообработки данных КМСС для определения высоты верхней границы облачности. Апробация рассмотренных алгоритмов стереообработки спектральнональных снимков выполнена по обзорным изображениям от КА «Ресурс-ДК», съемочная система которого аналогична КМСС и включает три смещенные в фокальной плоскости ПЗС-линейки. На рис. 3 приведен пример карты высот облачности, полученной по двум спектральнональным снимкам. Карта представляет собой изображение в картографической проекции, на котором присутствуют только облачные образования. Цвет облаков соответствует их высоте согласно легенде. Дальнейшая отработка разработанной технологии будет осуществлена на натурных данных в ходе летных испытаний КА «Метеор-М» №1.

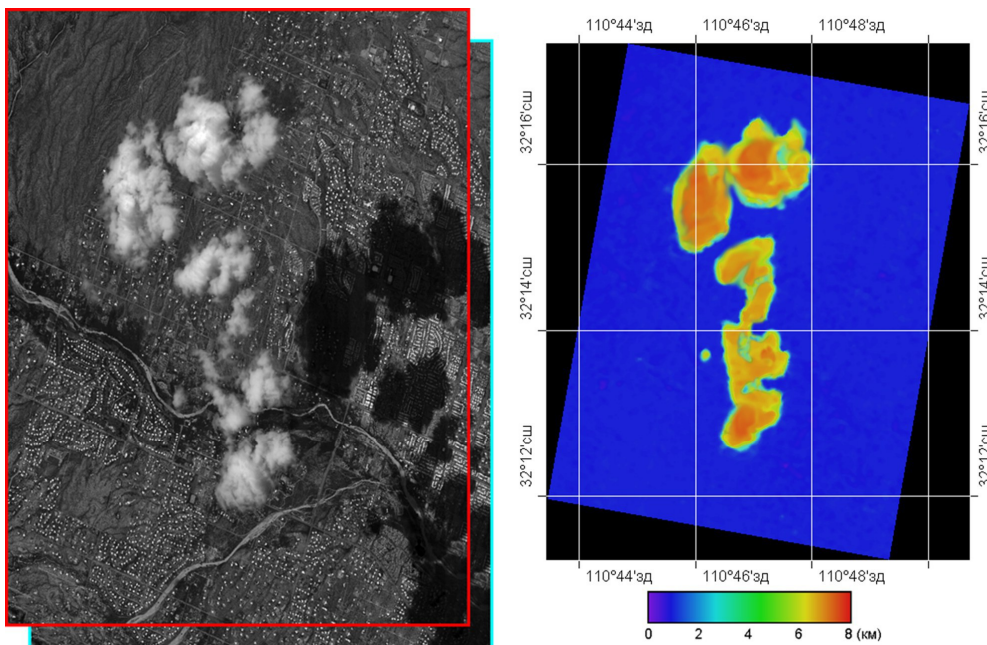


Рис. 3. Пример формирования карты высот облачности

Литература

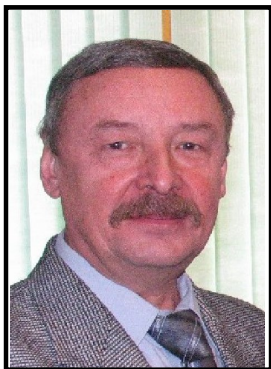
1. Nieman S.J., Schmetz J., Menzel W.P. A comparison of several techniques to assign heights to cloud tracers // J. Appl. Meteorol., 1993, 32, С.1559-1568.
2. Kahn B.H., Chahine M.T., Stephens G.L., Mace G.G., Marchand R.T., Wang Z., Barnet C.D., Eldering A., Holz R.E., Kuehn R.E., Vane D.G. Cloud type comparisons of AIRS, CloudSat, and CALIPSO cloud height and amount // Atmos. Chem. Phys., 2008, 8, С.1231-1248.
3. Fujita T. Principle of stereoscopic height computations and their applications to stratospheric cirrus over severe thunderstorms // Journal of the Meteorological Society of Japan, 1982, 60, С.355-368.
4. Акопов Г.А., Семенов В.Т., Чуркин А.Л. Задачи, решаемые космическим аппаратом «Метеор-М» по мониторингу Земли и околоземного космического пространства // Солнечно-земная физика, 2004, Вып.5. С.40-42.
5. Жуков Б.С., Василейский А.С., Железнов М.М., Жуков С.Б., Бекренев О.В., Пермитина Л.И. Задачи обработки мультиспектральных видеоданных КМСС на КА «Метеор-М» // Современные проблемы дистанционного зондирования Земли из космоса. М.: «ООО «Азбука-2000», 2007, Вып.4 Т.1, С.222-228.
6. Seiz G., Tjemkes S. Multiview Cloud-Top Height and Wind Retrieval with Photogrammetric Methods: Application to Meteosat-8 HRV Observations // Journal of Applied Meteorology and Climatology, 2007, 46(8), С.1182-1195.
7. Muller J.P., Denis M.A., Dundas R.D., Mitchell K.L., Naud C., Mannstein H. Stereo cloud-top heights and cloud fraction retrieval from ATSR-2 // International Journal of Remote Sensing, 2007, 28(9), С.1921-1938.
8. Naud C.M., Muller J.P., Intercomparison of multiple years of MODIS, MISR and radar cloud-top heights // Annales Geophysicae, 2005, 23(7), С.2415-2424.
9. Hirschmuller H. Accurate and efficient stereo processing by semi-global matching and mutual information // IEEE Conference on Computer Vision and Pattern Recognition, 2005, 2.
10. Owens J.D., Houston M., Luebke D., Green S., Stone J.E., Phillips J.C. GPU Computing // Proceedings of the IEEE, 2008, 96(5), С.879-899.
11. Buck I., Foley T., Horn D., Sugerman J., Fatahalian K., Houston M., Hanrahan P. Brook for GPUs: Stream Computing on Graphics Hardware // Proceedings of ACM SIGGRAPH, 2004.

CLOUD-TOP HEIGHT DETERMINATION FROM MULTISPECTRAL SCANNER IMAGING COMPLEX ONBOARD “METEOR-M”

Kuznetsov A.E., Poshehonov V.I.

Possibility of cloud-top height determination from images of multispectral scanner imaging complex onboard “Meteor-M” №1 is investigated. Technology and algorithms of fast stereo processing on modern hardware are considered.

ПАМЯТИ ПЕТРА ИВАНОВИЧА ЗУДКОВА



Ушел из жизни Петр Иванович Зудков, вице-президент Российского научно-технического общества радиотехники, электроники и связи им. А.С. Попова

П.И. Зудков родился 19 августа 1944 года в г. Москве. После окончания средней школы в 1961 году он поступил на физический факультет Московского государственного университета, который успешно закончил в 1967 году по специальности радиофизика.

По окончании Университета работал в НИИ «Полнос». Был главным разработчиком электрооптических затворов для лазерной техники специального применения. В 1977 году П.И. Зудков на VII съезде Всесоюзного научно-технического общества радиотехники, электроники и связи имени А.С. Попова был избран заместителем председателя Общества, а в 1991 году после реорганизации Всесоюзного Общества в Российское НТОРЭС им. А.С. Попова - вице-президентом Общества. В общей сложности более 30 лет П.И. Зудков отдал работе в Научно-техническом обществе радиотехники, электроники и связи им. А.С. Попова.

При активном участии П.И. Зудкова, продолжая традиции русских научно-технических обществ, Российское НТОРЭС им. А.С. Попова после трудных 90-х годов прошлого века вновь завоевывает авторитет среди ученых и специалистов как у нас в стране, так и за рубежом. Большой личный вклад П.И. Зудков внес в организацию и проведение, начиная с 1998 года, ежегодной международной конференции и выставки «Цифровая обработка сигналов и ее применение – DSPA» и издание научно-технического журнала «Цифровая обработка сигналов», выполняя обязанности ответственного редактора.

Душевность, отзывчивость и тактичность по отношению к окружающим снискали ему большое уважение коллег по работе.

Светлая память о нем навсегда сохранится в сердцах всех, кто знал этого замечательного человека, профессионала высокого класса.

Президиум Центрального совета РНТОРЭС им. А.С. Попова и редколлегия журнала «Цифровая обработка сигналов» выражают глубокое сочувствие семье П.И. Зудкова.

УДК 004.932

РЕАЛИЗАЦИЯ МНОГОФУНКЦИОНАЛЬНОГО КОДЕРА AVC

Гулевич А.Е., ФГУП «Главный радиочастотный центр», alex_gulevich@mail.ru

Дворкович А.В., д.т.н., ФГУП «Главный радиочастотный центр», a_dvork@nircom.ru

Кочарян А.Э., ФГУП «Главный радиочастотный центр», ashotkoch@mail.ru

Мингазов И.Д., ФГУП «Главный радиочастотный центр», elijacontact@gmail.com

Ключевые слова: кодер, стандарт, изображение, компрессия, компенсация движения, пиксел, предсказание, обработка, фильтрация, количество бит, видеотрансляция.

Современный стандарт видеокодирования H.264/AVC был разработан Совместной Видеогруппой (Joint Video Team), образованной Группой Экспертов по Видеокодированию МСЭ-Т (VCEG ITU-T) и Экспертной Группой по Движущимся Изображениям МСС (MPEG ISO/IEC). Он обозначается как Рекомендация H.264 МСЭ-Т и Стандарт ISO/IEC 14496-10 (MPEG-4 Part 10 AVC) в МСС (последняя редакция стандарта – март 2009г. [1]). Основными целями разработки стандарта H.264/AVC были заявлены повышенная эффективность компрессии и обеспечение удобного для транспортировки по различным сетям представления видео как для интерактивных, так и вещательных приложений [2]. Стандарт H.264/AVC существенно повысил эффективность видеокодирования по отношению к ранее принятым стандартам MPEG-2 Part 2 [3], MPEG-4 Part 2 [4], H.263 [5].

Повышение эффективности видеокодирования в стандарте H.264/AVC обеспечивается применением как уже известных методов и подходов, так и новых функциональных возможностей: компенсация движения с использованием переменных размеров блока, включая малые размеры блока; компенсация движения с точностью до четверти пиксела; вектора движения, выходящие за границы изображения; компенсация движения с несколькими опорными изображениями; независимость порядка воспроизведения изображений и порядка опорных изображений; независимость методов обработки изображений и возможности их использования для предсказания движения; взвешенное предсказание; улучшенная обработка «пропущенных» (skipped) блоков и блоков с «прямым» (direct) предсказанием движения; направленное пространственное предсказание для внутрикадрового кодирования; деблокировочная фильтрация в цикле кодирования; преобразование блоков небольшого размера; иерархическое преобразование блоков; преобразование с использованием 16-битной арифметики; точное обратное преобразование; арифметическое или контекстно-адаптивное энтропийное кодирование. В стандарте H.264/AVC повышена устойчивость к ошибкам и потерям данных, обеспечивается гибкость работы на множестве сетевых структур.

В статье рассмотрена реализация программного кодера стандарта H.264/AVC для систем вещания и двусторонней видеосвязи. Подробно рассмотрен алгоритм управления скоростью выходного потока кодера.

Существенным шагом вперед стала реализация методов кодирования с масштабированием (Scalable Video Coding) в этом стандарте. По сравнению с предшествующими стандартами значительно повышена эффективность кодирования и снижены требования к производительности декодера.

Авторами был разработан программный видеокoder H.264/AVC. Этот кодер использован в системах VPhone (система видеоконференцсвязи и вещания), AVIS (система узкополосного наземного цифрового мобильного аудио- и видеовещания). При разработке данных систем к ним предъявлялись требования работы с каналами ограниченной пропускной способности, низкой задержки, значительного сжатия при обеспечении максимально возможного качества видеoinформации. Кодер данного стандарта был выбран, как способный удовлетворить заданным требованиям.

В разработанном авторами программном кодере H.264/AVC реализованы основные важные возможности стандарта, использование которых, с одной стороны, существенно повышает качество кодирования, а с другой стороны не приводит к резкому росту вычислительных затрат. Кодер не поддерживает взвешенное предсказание, избыточные слои, разделение данных, High профиль, SP/SI синхронизацию/переключение изображений, кодирование с масштабированием. Разработанный кодер может функционировать в составе ТВ кодера стандартного разрешения реального времени, в составе системы видеотрансляций в Интернете, в составе систем видеоконференцсвязи.

Важнейшей частью кодера при использовании в различных приложениях является блок контроля потока. При фиксированном параметре квантования количество бит каждого закодированного макроблока изменяется в зависимости от содержания кадра, поэтому скорость потока на выходе кодера существенно неравномерна. Обычно кодер с постоянными параметрами производит больше бит, если в видеосюжете присутствует быстрое движение и/или мелкие детали, и меньше бит - в случае медленного движения и/или отсутствия мелких деталей.

Вариации скорости потока могут вызывать проблемы для многих практических задач доставки и хранения



видео. Например, канал постоянной пропускной способности не может передавать поток переменной скорости, если флуктуации скорости потока превышают пиковую пропускную способность канала. В таких случаях происходит либо потеря данных, либо накопление задержки. Таким образом, необходимо адаптировать или контролировать скорость потока, производимого видеокодером, для удовлетворения требований, накладываемых пропускной способностью канала и механизмом доставки. Другое ограничение связано с фиксированным объемом информационных носителей, поэтому для оптимального заполнения объема носителя также необходимо контролировать скорость потока закодированного видео.

Неравномерность скорости потока, производимого видеокодером, может быть сглажена буферизацией закодированных данных перед передачей. Поток переменной скорости, создаваемый кодирователем, помещается в FIFO-буфер, этот буфер освобождается с постоянной скоростью, которая равна пропускной способности канала (данные из буфера помещаются в канал). Другой FIFO-буфер размещается перед входом декодера, он заполняется данными из канала с постоянной скоростью (равной пропускной способности канала) и освобождается декодером с неравномерной битовой скоростью, но с постоянной кадровой скоростью. Таким образом, поток переменной скорости может быть адаптирован к постоянной скорости канала посредством использования буферов кодирователя и декодера. Но такая адаптация достигается ценой внесения задержки и использования памяти для буферов, и чем больше вариации скорости потока, тем требуются большие задержка и размеры буферов. При произвольных вариациях скорости потока этот метод невозможно использовать на практике, т.к. это может привести к недопустимо высоким задержкам и размерам буферов, а система передачи зачастую должна иметь обратную связь.

Контроль потока управляет параметрами кодирователя (шагом квантования) с целью получения заданной скорости потока и минимизации искажений декодированного видео при данной скорости. Достижение оптимального компромисса между скоростью потока и качеством - нетривиальная задача, здесь могут применяться различные подходы и алгоритмы, в зависимости от типа видеоприложения. Их можно классифицировать следующим образом:

- Кодирование не в реальном времени на носители. Время кодирования жестко не ограничено, поэтому могут применяться сложные алгоритмы. Целью является помещение закодированного видео на доступный объем носителя, при этом максимизируется качество и, возможно, учитывается, что буфер декодирующего устройства или программы не должен переполняться или опустошаться в процессе декодирования. Могут применяться двухпроходные схемы (во время первого прохода кодирователь собирает статистику о видеопоследовательности, кодирование производится во время второго прохода).
- Кодирование живого видео для вещания. Процесс декодирования и буферизации может иметь ограни-

чения, в то время как кодирование может использовать мощное оборудование. Обычно допустима задержка в несколько секунд. Здесь применяются алгоритмы контроля потока средней сложности, возможно, включая двухпроходное кодирование каждого кадра.

- Кодирование для двусторонней видеоконференции. Каждый терминал производит как кодирование, так и декодирование, и вычислительная мощность может быть ограничена. Задержка должна быть минимальной (обычно в пределах 0,1 – 0,8 секунды). Здесь обычно применяются алгоритмы контроля потока невысокой сложности. Размеры буферов кодирователя и декодера должны быть минимизированы для обеспечения минимальной задержки, кодирователь должен точно выдерживать скорость выходного потока. Качество декодированного видео может значительно варьироваться, например, возможно существенное ухудшение качества при быстром движении, смене сюжета и т.п. Однако для некоторых задач (видеонаблюдение, телеконсультации) важно сохранять достаточное высокое качество изображения, при этом возможно пожертвовать постоянной кадровой скоростью. В этом случае применяется механизм пропуска кадров, он позволяет удовлетворить одновременно требования низкой задержки и постоянной, достаточно высокой четкости изображения.

В разработанном кодере H.264/AVC реализованы все перечисленные варианты регулирования выходного потока.

Алгоритм контроля потока разработанного кодирователя H.264/AVC построен на базе схемы SRC (Scalable Rate Control) [6-8]. Алгоритм выдерживает заданную скорость потока на сегменте кадров. Он использует следующую модель для скорости потока:

$$R = S \cdot \left(\frac{X_1}{Q} + \frac{X_2}{Q^2} \right), \quad (1)$$

где R – скорость потока, Q – шаг квантования, S – мера сложности кадра (в данной реализации – средняя абсолютная разность остаточного кадра после компенсации движения), X_1 и X_2 – параметры модели. Алгоритм состоит из следующих шагов, которые выполняются после компенсации движения для i -того кадра:

1. Вычисление целевой скорости потока R_i , основываясь на количестве кадров в сегменте кадров, числе бит, доступных для остатка сегмента кадров, и оценочной сложности i -го кадра (в качестве оценки используется размер предыдущего кадра). Если предыдущий кадр имел высокую сложность, то предполагается, что следующий кадр тоже будет сложным и должен занять большое число бит. Алгоритм пытается найти баланс между этим требованием и лимитом на число бит для сегмента:

$$R_{i+1} = \frac{G_i}{N_i} \cdot (1 - W) + A_i \cdot W, \quad (2)$$

где R_{i+1} – целевая скорость потока для $i+1$ кадра, G_i – число бит, доступных для остатка сегмента кадров на момент времени i , N_i – число кадров в остатке сегмента кадров на момент времени i , A_i – число бит, использованных для кодирования i -того кадра, W – взвешивающий параметр.

Далее учитываются требования размера буфера:

$$\text{Если } B_{i+1} + R_{i+1} > 0.9 \cdot V, \text{ то } R_{i+1} = 0.9 \cdot V - B_{i+1}, \quad (3)$$

где B_i – заполненность буфера, V – размер буфера. Цель данного выражения – снизить вероятность переполнения буфера и выхода за пределы заданной задержки.

2. Вычисление шага квантования Q_i (вычисляется S для всего остаточного кадра и решается уравнение 1).

3. Кодирование кадра.

3. Обновление параметров модели X_1, X_2 . Используется количество бит, фактически затраченных на кодирование i -того кадра R_i . Для этого выбираются точки данных, используя окно, размер которого зависит от изменения сложности видео. Если сложность изменяется значительно, то используется небольшое окно с предыдущими данными. При небольшом изменении сложности используется больше предыдущих данных. Далее методом линейной регрессии вычисляются параметры модели X_1, X_2 :

$$X_2 = \frac{n \sum_{i=1}^n R_i - (\sum_{i=1}^n Q_i^{-1})(\sum_{i=1}^n Q_i R_i)}{n \sum_{i=1}^n Q_i^{-2} - (\sum_{i=1}^n Q_i^{-1})^2},$$

$$X_1 = \frac{(\sum_{i=1}^n Q_i R_i) - X_2 \sum_{i=1}^n Q_i^{-1}}{n}. \quad (4)$$

После вычисления новых параметров из рассмотрения выбрасываются статистически «плохие» точки, для которых разность теоретически и практически затраченных бит превосходит величину стандартного отклонения. Таким образом, получается новый, более репре-

зентативный набор данных, для которого еще раз пересчитываются параметры модели по формуле 4.

Результаты работы приведенного алгоритма проиллюстрированы на рис. 1. Была использована последовательность кадров из фильма «Матрица» размера CIF, 25 кадров/сек, задержка кодирования не более 0,25 сек, выходной поток 300 кбит/сек. По горизонтали отложены номера кадров, размеры кадров – в байтах, задержка – в секундах. Из рис. 1 видно, что в районе 370-го кадра (смена сюжета) задержка резко выросла (более 3 сек), то есть механизм управления скоростью потока обеспечивает нужный поток, но не обеспечивает задержку.

Для предотвращения переполнения буфера изменяется механизм пропуска кадров: если кодер предсказывает, что следующий кадр вызовет переполнение буфера, то следующий кадр пропускается. Это эффективный метод защиты от переполнения буфера, но он может вызвать значительное снижения качества восприятия видео в случае пропуска нескольких последовательных кадров. Для борьбы с этой проблемой алгоритм пропуска кадров несколько модифицируется и выглядит следующим образом. Перед кодированием следующего кадра, кодер вычисляет заполненность буфера и оценивает размер следующего кадра (используется размер предыдущего). Если их сумма выше некоторого порога (например, 80% от размера буфера), то следующий кадр пропускается. Данный порог применяется для снижения вероятности пропуска группы последовательных кадров. Он может быть адаптивным или являться некоторой предопределенной константой.

На рис. 2 проиллюстрирована работа такого модифицированного алгоритма. Теперь заданная задержка выдерживается.

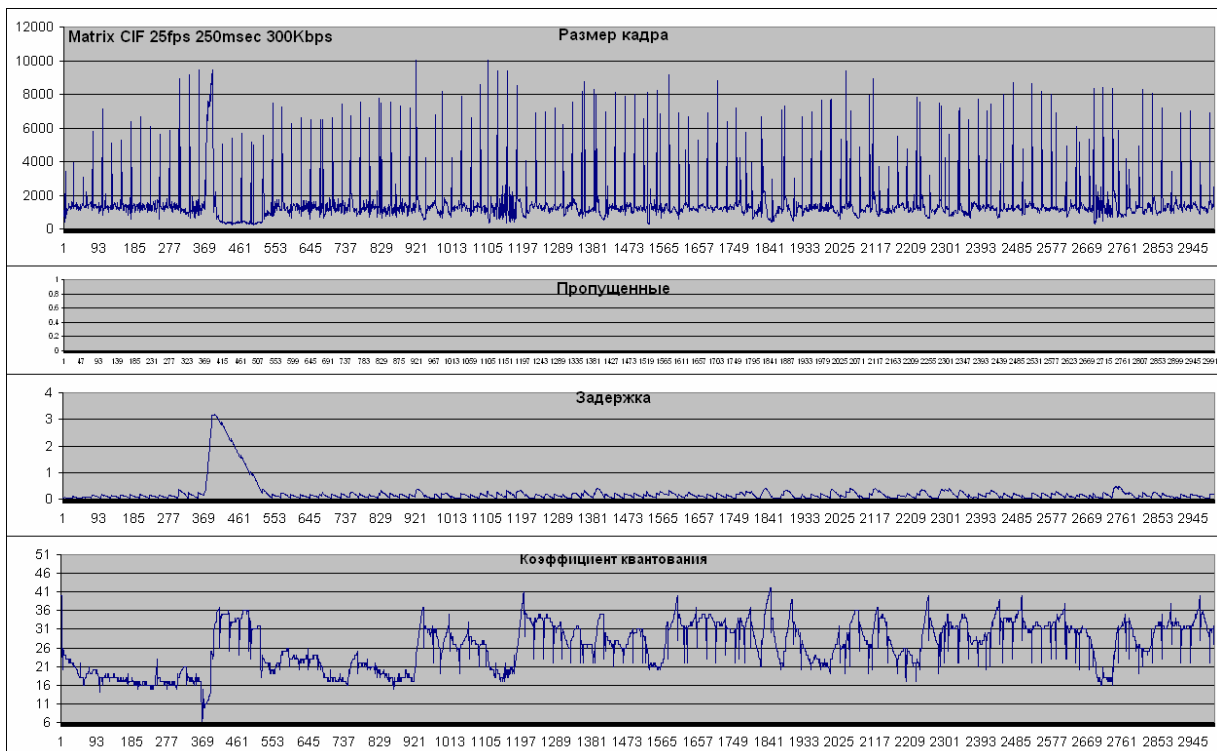


Рис. 1. Кодирование видеопоследовательности без пропуска кадров.

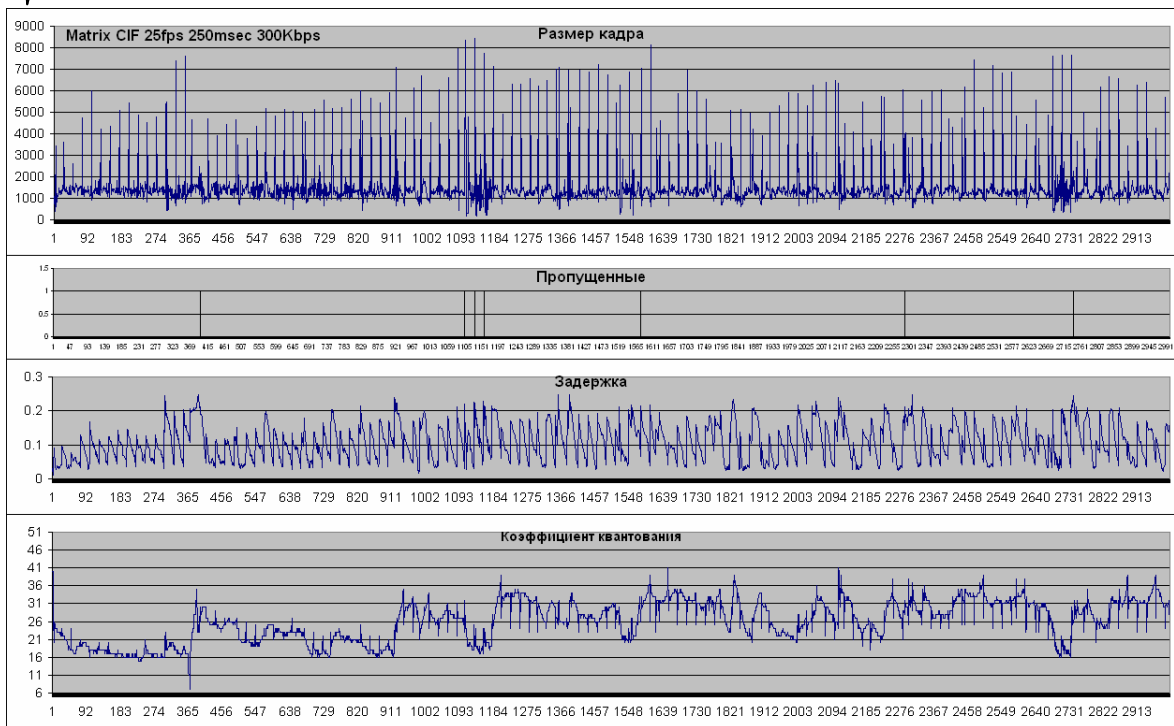


Рис. 2. Кодирование видеопоследовательности с пропуском кадров.



Рис. 3. Ухудшение качества I-кадра при фиксированном соотношении размеров I- и P-кадров (второй кадр – типа I).

Еще одной проблемой является неравномерность качества I- и P-кадров (то есть кадров, закодированных без предсказания движения и с предсказанием движения). Изначально в кодере был реализован вариант, когда в параметрах кодера задается соотношение размеров I- и P-кадров. Проблемой данного подхода является то, что низкое отношение вызывает визуально заметное ухудшение качества I-кадров (см. рис. 3), а высокое – увеличивает вероятность переполнения буфера. Кроме того, для разных пространственных разрешений одно отношение размеров I- и P-кадров может создавать раз-

ные соотношения качества I- и P-кадров.

Поэтому был реализован следующий подход: для I-кадров выбор шаг квантования вычисляется, как коэффициент квантования предыдущего P-кадра, деленный на константу $r = 1.4$ (пункты 1-2 пересматриваются). Данная константа выбрана эмпирически и позволяет получать I- и P-кадры примерно одинакового качества. При этом учитывается требование размера буфера: из уравнения (1) и условия (3) можно получить минимальный допустимый шаг квантования. Этот подход проиллюстрирован на рис. 4.



Рис. 4. Отсутствие ухудшения качества I-кадра при фиксированном соотношении коэффициентов квантования I- и P-кадров (второй кадр – типа I).

Дальнейшим улучшением алгоритма является перекодирование кадра, если реальный размер закодированного кадра сильно отличается от целевого. Это поможет улучшить качество кодирования и снизить количество пропускаемых кадров.

Данная схема контроля потока и пропуска кадров применима для различных битовых скоростей и временных/пространственных разрешений. Описанный алгоритм контроля потока используется в видеоинформационной системе VPhone. Система обеспечивает высокое качество видеоизображения и малые задержки при интерактивном общении. Достоинства системы особенно проявляются на низкоскоростных каналах связи и каналах связи с жестким ограничением скорости передачи (например, спутниковый канал).

Литература

1. ITU-T Recommendation H.264. Advanced Video Coding for Generic Audiovisual Services (ISO/IEC 14496-10. Information Technology – Coding of Audio-Visual Objects – Part 10: Advanced Video Coding). - ITU-T, 2009.
2. Дворкович А.В. Эффективное кодирование видеоинформации в новом стандарте H.264/AVC // Труды НИИП. - М., 2005. - С. 56-69.
3. ITU-T Recommendation H.262, ISO/IEC 13818-2. Information Technology – Generic Coding of Moving Pictures and Associated Audio Information: Video (MPEG-2 Video). - ITU-T, 2000.
4. ISO/IEC 14496-2 – Information Technology – Coding of Audio-Visual Objects – Part 2: Visual (MPEG-4 Video). - ISO, 2004.
5. ITU-T Recommendation H.263. Video Coding for Low Bit Rate Communication. - ITU-T, 2005.
6. H.-J. Lee, T. Chiang, Y.-Q. Zhang. Scalable rate control for MPEG-4 video. // IEEE Trans. on Circuits and Systems for Video Technol. – Sep. 2000, v. 10, № 6, pp. 878-894.
7. T. Chiang, Y.-Q. Zhang. A new rate control scheme using a new rate-distortion model. // IEEE Trans. on Circuits and Systems for Video Technol. – Feb. 1997, v. 7, № 1, pp. 246–250.
8. Дворкович А.В., Быстров С.В. Методика регулировки потока при сжатии динамических видеоизображений // Цифровая обработка сигналов. - 2003. - № 2. - С. 12-16.

АВТОМАТИЧЕСКОЕ СЛЕЖЕНИЕ ЗА ОБЪЕКТАМИ ПРИ ОТСУТСТВИИ АПРИОРНЫХ СВЕДЕНИЙ О ФОНОЦЕЛЕВОЙ ОБСТАНОВКЕ

Алпатов Б.А., д.т.н., зав. кафедрой АИТУ Рязанского государственного радиотехнического университета (РГРТУ)

Бабаян П.В., к.т.н., начальник Научно-исследовательской лаборатории автономных информационно-управляющих систем РГРТУ, aitu@rgta.ryazan.ru

Смирнов С.А., м.н.с. кафедры АИТУ РГРТУ

Ключевые слова: автоматическое слежение, последовательность изображений, система признаков, видеонаблюдение, алгоритм сопровождения, объект, фоноцелевая обстановка.

Введение

Алгоритмы автоматического слежения за объектами по данным видеонаблюдений имеют широкий спектр применения в системах оптической локации, управления движением воздушных и наземных транспортных средств, в робототехнических системах, в бортовых системах навигации пилотируемой и беспилотной техники, в комплексах аэрокосмического мониторинга Земли, поиска и спасения людей в чрезвычайных ситуациях.

Разработано большое количество алгоритмов сопровождения объектов, каждый из которых характеризуется определёнными требованиями к типам фоноцелевой обстановки, при которых алгоритм работоспособен [1]. Здесь и далее под фоноцелевой обстановкой будем понимать совокупность условий наблюдения, характеризующихся заданными статическими и динамическими свойствами объектов, фона, освещения, атмосферы, устройств формирования и передачи изображений.

Во многих случаях, имея априорные сведения о фоноцелевой обстановке, можно выбрать алгоритм слежения за объектами, обеспечивающий высокие качественные показатели слежения. Однако, достаточно часто, особенно при работе автономных комплексов анализа изображений реального времени, априорные сведения о фоноцелевой обстановке отсутствуют, так же как и отсутствует возможность оперативного вмешательства оператора для задания типа фоноцелевой обстановки. В данной ситуации необходимо использовать алгоритмы слежения за объектами, работоспособные в широком диапазоне условий наблюдения. К сожалению, на сегодняшний момент среди существующих методов слежения нет такого, который обладал бы удовлетворительным качеством в любых условиях наблюдения. Это обстоятельство заставляет искать решение задачи в оценивании характеристик фоноцелевой обстановки и использовании этих характеристик для автоматического выбора того или иного алгоритма анализа изображений. Один из таких подходов предложен в настоящей работе.

ПОСТАНОВКА ЗАДАЧИ. Будем описывать наблюдаемые видеосюжеты функциями трёх переменных $l(i, j, n)$, где (i, j) – координаты точки изображения, n – номер кад-

Рассмотрена проблема слежения за объектами в системах анализа последовательностей изображений при отсутствии априорной информации об условиях наблюдения. Предложено решение проблемы путём автоматического выбора наиболее эффективного метода слежения на основе анализа ряда признаков наблюдаемой видеопоследовательности. Представлены результаты статистических исследований.

ра. В настоящей работе рассматриваются следующие основные типы фоноцелевой обстановки, соответствующие наблюдаемым видеосюжетам:

- автомобили, наблюдаемые на фоне леса, травы, домов, дорог, городского ландшафта;
- самолёты, вертолёты, наблюдаемые на фоне облачного и ясного неба;
- морские суда, наблюдаемые на фоне водной поверхности или неба вблизи линии горизонта.

Объекты на данных видеопоследовательностях наблюдаются как на относительно однородных фонах, так и на сложных неоднородных фонах с малым отношением сигнал/шум. Фон предполагается неподвижным, объекты могут быть подвижными или неподвижными. Сюжеты односпектральные, спектральный диапазон наблюдения – видимый или инфракрасный. Размер объектов – от 4x4 до 60x60 точек. Кроме этого, на наблюдаемых изображениях присутствует шум, порождаемый датчиком изображения, а также шум квантования изображения по уровню, вызываемый ограниченной разрядностью представления значений яркости.

Конкретный тип фоноцелевой обстановки, соответствующей текущему наблюдаемому видеосюжету, неизменен во времени, но заранее неизвестен.

На кадре с номером n_0 оператор осуществляет целеуказание, то есть задаёт пару координат (i_0, j_0) , соответствующую центру объекта, за которым необходимо осуществлять слежение.

Дальнейшее слежение может осуществляться одним из четырёх алгоритмов, которые известны из литературы и в настоящей работе подробно не рассматриваются:

- алгоритм на основе корреляционного совмещения изображений со сглаживанием эталона (алгоритм I) [2];
- алгоритм на основе статистической сегментации (алгоритм II) [3];
- алгоритм на основе пространственной фильтрации

(алгоритм III) [4];

– алгоритм на основе пространственно-временной фильтрации (алгоритм IV) [1,5].

Задача состоит в том, чтобы, используя изображения из наблюдаемой видеопоследовательности при $n \leq n_0$, а также начальное целеуказание оператора, принять решение о том, какой из перечисленных выше четырёх алгоритмов использовать для продолжения слежения за объектом. При этом должно обеспечиваться наилучшее качество слежения.

Общий подход к решению задачи

Для решения поставленной выше задачи прежде всего необходимо конкретизировать критерий, по которому будут происходить сравнение алгоритмов с точки зрения качества слежения. Охарактеризуем некоторые известные из литературы подходы к решению задачи оценивания качества слежения. Оценивание качества слежения связано с анализом количественных метрик, описывающих эффективность слежения. Можно выделить два типа таких метрик.

Первый тип предполагает использование эталонной информации о траектории движения объекта, его форме и размерах. В частности, в работе [6] предлагается использовать такие метрики, как доля успешно отслеживаемых объектов; время завязывания траектории; количество потерь объектов за заданный промежуток времени; отношение числа ложных траекторий к общему их числу, среднеквадратическая ошибка слежения и другие. Косвенно оценить качество слежения можно с помощью критериев качества сегментации, таких как частоты правильного и ложного выделения точек объекта [7,8]. В работе [8] основными мерами качества являются коэффициент повышения отношения сигнал/фон, коэффициент подавления фона, коэффициент улучшения пикового отношения сигнал/фон.

Второй тип метрик не предполагает использования эталонных данных. Метрики вычисляются на основе информации, получаемой в процессе обработки наблюдаемой видеопоследовательности. Часто при вычислении метрик второго типа выполняются предварительные операции грубой сегментации изображения или оценивания координат объекта. Результат предварительных операций служит основой для расчёта метрик. В частности, в работе [9] приведено исследование девяти метрик, среди которых наиболее эффективными себя показали разность энтропий объекта и фона, статистика Колмогорова-Смирнова в комбинации с площадью объекта. В [7] рассмотрены метрики, основанные на анализе однородности, контрастности и формы объектов. В [10] качество слежения оценивается по значениям апостериорной плотности вероятности.

В дальнейшем метрики первого типа будем называть количественными характеристиками работоспособности (КХР), а метрики второго типа – признаками работоспособности (ПР). Преимуществом КХР является то, что они позволяют сравнивать работоспособность различных алгоритмов и, таким образом, объективно выбирать наилучший алгоритм. Недостаток КХР заключается в необходимости иметь эталонные данные. Для вычисления ПР не требуются эталонные данные, однако ПР дают лишь оценочную, не всегда адекватную характеристику работоспособности. Ещё более существенным недостатком ПР по сравнению с КХР является невозможность непосредственного сравнения качества слежения для различных алгоритмов, поскольку ПР, характеризующие разные алгоритмы слежения, обычно не совпадают ни по размерности, ни по диапазону значений.

Несмотря на различия между сущностями ПР и КХР, между ними существует прочная статистическая связь. Исследование этой статистической связи применительно к некоторым типам КХР и ПР приведено в [9].

В известных работах, на которые мы сослались, в основном затрагиваются проблемы выбора единственного алгоритма для использования в условиях заданной фоноцелевой обстановки. В настоящей работе основное внимание уделяется разработке методики автоматического выбора алгоритма в зависимости от текущей фоноцелевой обстановки.

Предлагаемая методика заключается в выполнении следующих этапов:

1. В момент целеуказания вычисляются ПР для каждого из алгоритмов. При вычислении ПР используется результат предварительного обнаружения движущихся объектов. Предварительное обнаружение осуществляется алгоритмом на основе пространственно-временной фильтрации.

2. На основании найденной заранее связи между вычисленными ПР и КХР определяются КХР для каждого из алгоритмов.

3. На основе сравнения КХР производится выбор алгоритма слежения, обеспечивающего наилучшее значение КХР.

Структура системы слежения за объектами, в которой может быть реализована предлагаемая методика, приведена на рис. 1. Отметим, что алгоритм на основе пространственно-временной фильтрации работает постоянно. После выполнения целеуказания начинается слежение за объектом одним из алгоритмов, либо продолжается слежение алгоритмом на основе пространственно-временной фильтрации.

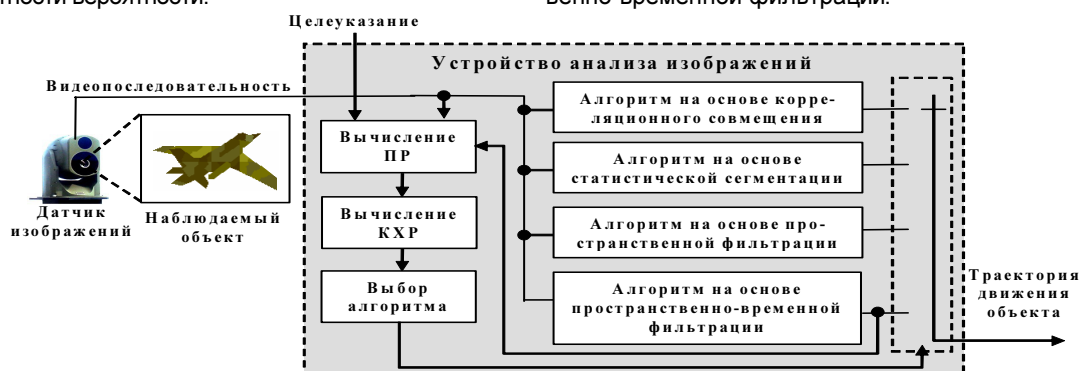
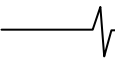


Рис. 1 – Структурная схема системы автоматического слежения за объектами



Для реализации представленной методики необходимо решить следующие вопросы:

1. Выбор КХР.
2. Выбор ПР для каждого алгоритма.
3. Нахождение статистической связи между ПР и КХР.
4. Разработка правила принятия решения о выборе того или иного алгоритма слежения.

Рассмотрим эти вопросы подробнее.

Выбор количественных характеристик работоспособности

Чаще всего КХР описываются скалярными величинами, принимающими значения из непрерывного множества или множества натуральных чисел. В настоящей работе предлагается использовать двоичную КХР, принимающую значения «следит» (1) и «не следит» (0). Целесообразность введения двоичной КХР будет обоснована ниже при рассмотрении вопроса о нахождении статистической связи между КХР и ПР. Остановимся на критерии успешного слежения: если в течение 70% кадров наблюдаемой видеопоследовательности измеренный центр объекта не выходит за пределы окружности, центр которой соответствует эталонному центру объекта, а радиус равен $\sqrt{a^2 + b^2} / 2$ (где a, b – эталонные размеры объекта), то КХР имеет значение «следит» (1), в противном случае критерий имеет значение «не следит» (0).

ВЫБОР ПР. Каждый ПР должен служить индикатором качества слежения для соответствующего алгоритма. Очевидно, что сущность ПР должна быть связана с теми особенностями объекта, которые тот или иной алгоритм слежения использует для отделения объекта от фона. Поскольку ни один из рассматриваемых алгоритмов слежения не использует информацию о конкретных значениях яркости объекта и фона и о характере геометрических преобразований, то необходимо обеспечить инвариантность ПР к изменению яркости, контраста, масштаба, к вращению и сдвигу изображения.

Перед вычислением ПР должна быть произведена предварительная оценка прямоугольного множества точек, принадлежащих объекту. Обозначим эту оценку через H . Множество H определяется на основе целеуказания оператора и результатов предварительного обнаружения движущихся объектов. Результат предварительного обнаружения объектов представляет собой множество прямоугольных областей, соответствующих обнаруженным объектам. Если целеуказание (i_0, j_0) попадает в какую-либо прямоугольную область, то эта область принимается за множество H . Если целеуказание не попадает ни в одну прямоугольную область, то (i_0, j_0) принимается за центр множества H , а размеры области H принимаются равными некоторому заданному заранее значению.

Для рассматриваемых алгоритмов слежения предлагается использовать следующие ПР.

ПР, характеризующий методы корреляционно-совмещения, основан на сравнении временной изменчивости объекта и степени отличия объекта от фона. Если на кадре n_0 было выполнено целеуказание, то

за эталонное изображение объекта принимается изображение в области H . Тогда ПР можно определить следующим образом:

$$KF_{кор} = \frac{F_g}{F_h}, \quad (1)$$

где F_g – оценка степени различия эталонного изображения объекта и фона, F_h – оценка межкадровой изменчивости изображения объекта. F_g определяется как минимальное значение разностной критериальной функции в области, граничащей с объектом на кадре n_0 . F_h определяется как минимальное значение критериальной функции на кадре $n_0 + 1$.

Если межкадровая изменчивость объекта F_h больше степени отличия эталонного изображения объекта и фона F_g , следует ожидать уверенного сопровождения объекта.

ПР, характеризующий алгоритм на основе статистической сегментации, основан на вычислении меры статистического различия объекта и фона. Если $\hat{P}(X|O)$ – гистограмма яркости изображения в области H , ограничивающей объект, $\hat{P}(X|\Phi)$ – гистограмма яркости изображения в окрестности множества H , то меру статистического различия объекта и фона можно вычислить при помощи критерия Джинни [11]:

$$KF_{сегм} = \sum_x |\hat{P}(X|O) - \hat{P}(X|\Phi)|. \quad (2)$$

ПР, характеризующий алгоритм на основе пространственной фильтрации, основан на оценке амплитуды яркостного импульса, представляющего объект на изображении:

$$KF_{прост} = \frac{\max_{(i,j) \in H} |l(i,j,n_0) - \hat{m}_g|}{\hat{\sigma}_g}, \quad (3)$$

где $l(i,j,n_0)$ – яркость изображения в точке (i,j) на кадре n_0 , \hat{m}_g – оценка средней яркости фона, $\hat{\sigma}_g$ – оценка СКО фона. Оценки \hat{m}_g и $\hat{\sigma}_g$ рассчитываются по окрестности множества H . Отметим, что числитель выражения (3) представляет собой максимальный контраст объекта.

ПР, характеризующий алгоритм на основе пространственно-временной фильтрации, основан на оценке степени отличия объекта от фона:

$$KF_{прост.-времен.} = \frac{1}{M \cdot N} \sum_{(i,j) \in H} \frac{|\hat{g}(i,j,n_0) - l(i,j,n_0)|}{\hat{\sigma}(i,j,n_0)}, \quad (4)$$

где $\hat{g}(i,j,n_0)$ – оценка яркости фонового изображения в точке (i,j) , $\hat{\sigma}(i,j,n_0)$ – оценка СКО фона, M, N – размеры объекта. Оценки $\hat{g}(i,j,n_0)$ и $\hat{\sigma}(i,j,n_0)$ определяются в процессе работы алгоритма слежения. Соответствующие процедуры оценивания описаны в [1].

Нахождение статистической связи между признаками и количественными характеристиками работоспособности

Ввиду чрезвычайной сложности построения математических моделей видеосюжетов, практически отсутствует возможность аналитического построения функции,

устанавливающей соответствие между значениями ПР и значениями КХР. Поэтому в настоящей работе данная задача решается статистически с применением базы данных видеопоследовательностей с разной фоноцелевой обстановкой.

Используемая база данных содержала 50 сюжетов. Продолжительность сюжетов составляла от 200 до 400 кадров. Фоноцелевая обстановка соответствовала описанной в разделе «постановки задачи».

Предварительно для каждого видеосюжета человеком-экспертом были сформированы эталонные данные, описывающие «точные» координаты и размеры объекта. Точность нахождения связи между ПР и КХР зависит от объёма базы данных, разнообразия сюжетов, участвующих в статистической обработке, а также от количества значений, которые может принимать КХР. В рассматриваемом случае объём базы данных является недостаточным для формирования функции зависимости КХР от значений ПР при большом количестве возможных значений КХР. Именно по этой причине принято решение ограничиться двоичной КХР. Переход от значений ПР к двоичной КХР должен осуществляться путём сравнения признаков работоспособности с пороговыми значениями.

Для получения пороговых значений была исследована работоспособность алгоритмов для каждой из видеопоследовательностей. Решение о работоспособности того или иного алгоритма слежения принималось на основе сравнения эталонных данных с результатами измерения координат и вычисления КХР.

Рассмотрим процедуру нахождения порогового уровня для произвольного алгоритма слежения. Всё множество видеосюжетов можно разбить на подмножество X_1 , для которых КХР приняла значение 1, и подмножество X_2 , включающее все остальные видеосюжеты.

В качестве критерия выбора порога будем использовать критерий максимального правдоподобия. Для этого приближенно опишем плотность распределения ПР внутри каждого множества X_1 и X_2 с помощью нормального распределения и найдём выборочные средние m_i и дисперсию σ_i для каждого множества X_i . Поиск порогового значения сводится к вычислению точки пересечения гауссоид по известным соотношениям

$$\alpha_{1,2} = \frac{-(2m_1\sigma_2^2 - 2m_2\sigma_1^2) \pm \sqrt{D}}{2(\sigma_1^2 - \sigma_2^2)}, \quad (5)$$

где

$$D = (2m_1\sigma_2^2 - 2m_2\sigma_1^2)^2 - 4(\sigma_1^2 - \sigma_2^2) \cdot (\sigma_1^2 m_2^2 - \sigma_2^2 m_1^2 - 2\sigma_1^2 \sigma_2^2 \log \sqrt{\sigma_1 / \sigma_2})$$

Из полученных $\alpha_{1,2}$ выберем в качестве порогового значения то, которое лежит в диапазоне $[m_1, m_2]$.

В таблице 1 приведены пороговые значения для каждого из четырех алгоритмов.

Выбор алгоритма слежения

После перехода от значений ПР к значениям КХР необходимо принять решение, какой алгоритм слежения применять в текущих условиях наблюдения.

Таблица 1.

Пороговые значения

Алгоритм	Пороговое значение
Корреляционное совмещение	1,4
Статистическая сегментация	0,7
Пространственная фильтрация	6
Пространственно-временная фильтрация	1,3

В случае, если единичное значение ПР присутствует лишь для одного алгоритма слежения, то выбор алгоритма для данного видеосюжета соответствует единичному ПР. В противном случае предлагается производить выбор алгоритма, обладающего самым высоким приоритетом. Перечень алгоритмов в порядке убывания приоритета следующий: алгоритм на основе статистической сегментации, алгоритм на основе пространственно-временной фильтрации, алгоритм на основе пространственной фильтрации, алгоритм на основе корреляционного совмещения изображений.

Приведенный порядок выбора алгоритмов слежения можно объяснить следующими фактами. Алгоритм слежения должен быть устойчив к геометрическим изменениям объекта, вызванных его движением. С этой точки зрения алгоритм на основе корреляционного совмещения изображений является наиболее уязвимым. Приоритет алгоритма на основе статистической сегментации над алгоритмом пространственно-временной фильтрации связан с его инвариантностью к геометрическим преобразованиям фона. С другой стороны, алгоритм статистической сегментации более устойчив при слежении на неоднородном фоне, чем алгоритм на основе пространственной фильтрации.

В случае, когда значения всех ПР равны нулю, принимается решение о невозможности слежения.

Экспериментальные исследования

Экспериментальная проверка, проведённая на реальных видеосюжетах, показала высокую эффективность предлагаемой методики. На рис. 2 приведены примеры кадров из нескольких видеосюжетов, участвующих в экспериментальных исследованиях. Прямоугольниками отмечены положения объектов на кадрах. Линиями показаны траектории движения объектов. Крестами отмечены координаты объекта через 15 кадров.

В ходе принятия решения возникали ошибки двух видов:

- ошибка первого рода – ошибочное принятие решения, что алгоритм неработоспособен,
- ошибка второго рода – ошибочное принятие решения, что алгоритм работоспособен.

В таблице 2 приведены значения ошибок классификации для каждого алгоритма.

Полученные значения ошибок в большинстве случаев являются приемлемыми для использования предложенной методики выбора метода слежения. Возможной доработки, очевидно, требует признак для алгоритма статистической сегментации.

Ошибки классификации для различных алгоритмов

Алгоритм	Ошибка первого рода, %	Ошибка второго рода, %
Корреляционное совмещение (I)	0	0
Статистическая сегментация (II)	15,38	32,26
Пространственная фильтрация (III)	5,56	8,7
Пространственно-временная фильтрация (IV)	0	10,81

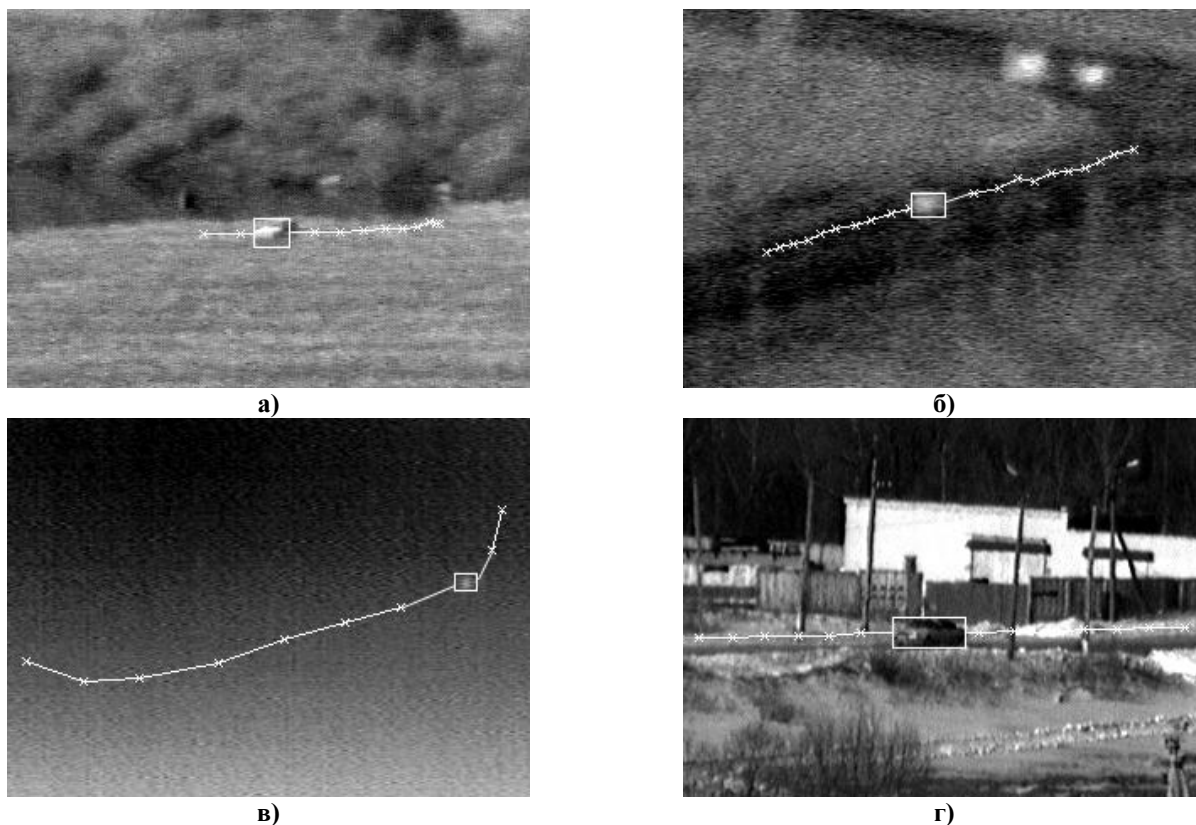


Рис. 2 – Кадры из нескольких видеосюжетов, участвующих в экспериментальных исследованиях.

На видеосюжете приняты решения об использовании следующих алгоритмов:

а) алгоритм I, б) алгоритм II, в) алгоритм III, г) алгоритм IV

Заключение

Результаты экспериментальных исследований позволяют сделать выводы о применимости предложенной методики выбора метода слежения при неизвестной фоноцелевой обстановке. Невысокая вычислительная сложность позволяет рекомендовать методику для использования в бортовых системах обнаружения и сопровождения объектов. В ходе дальнейших исследований предполагается расширить перечень возможных алгоритмов слежения, а также перейти к принятию решения о выборе алгоритма на каждом кадре.

Работа проводилась при финансовой поддержке Федерального агентства по науке и инновациям (госконтракт № 02.740.11.0002).

Литература

1. Методы автоматического обнаружения и сопровождения объектов. Обработка изображений и управление / Алпатов Б.А., Баба-ян П.В., Балашов О.Е., Степашкин А.И. – М.: Радиотехника, 2008. – 176 с.
2. Алпатов Б. А., Селяев А. А., Степашкин А.И. Цифровая обработка изображений в задаче отслеживания движущегося объекта // Изв. вузов. Сер. Приборостроение. – 1985. – №2. – С. 39-43.
3. Денисов Д.А., Низовкин В.А. Сегментация изображений на ЭВМ // Зарубежная радиоэлектроника. –1985. – №10 – С. 5-30.

4. Муравьев В.С., Муравьев С.И. Алгоритм выделения и измерения координат объектов, наблюдаемых на облачных фонах // Вестник РГРТУ, Рязань. – 2007. – №21. – С. 20-24.
5. Алпатов Б. А., Блохин А. Н. Модели и алгоритмы обнаружения и выделения движущихся фрагментов изображений // Автометрия. – 1995. – №4. – С. 100-104.
6. Ronald Rothrock, Oliver Drummond Performance metrics for multiple-sensor multiple-target tracking // Proc. of SPIE Vol. 4048, Signal and Data Processing of Small Targets 2000, pp. 521-531.
7. Yu-Jin Zhang Advances in image and video segmentation IRM Press 2006
8. Courtney I. Hilliard Selection of a Clutter Rejection Algorithm for Real-Time Target Detection from an Airborne Platform // Proc. of SPIE Vol. 4048, Signal and Data Processing of Small Targets 2000, pp. 74-84
9. Lloyd G. Clark, Vincent J. Veltin Image characterization for automatic target recognition algorithm evaluations // Proc. of SPIE Vol. 1310, Signal and Image Processing Systems Performance Evaluation, 1990, pp. 68-78.
10. Baoxin Li, Rama Chellappa A Generic Approach to Simultaneous Tracking and Verification in Video // IEEE Transactions on Image Processing. Vol. 11. No. 5. May 2002
11. Кобзарь А.И., Прикладная математическая статистика. Для инженеров и научных работников. – М.: ФИЗМАТЛИТ, 2006. – 816 с.

Посвящается юбилею журнала «Цифровая Обработка Сигналов»

ОСНОВНОЙ ВИДЕОПРОЦЕССОР НАШЕЙ ЦИВИЛИЗАЦИИ – УНИКАЛЬНЫЙ ПРИМЕР ЭВОЛЮЦИИ В НОВЕЙШЕЙ ИСТОРИИ?

Киричников А.П., член редколлегии журнала ЦОС

Введение

Данная работа преднамеренно выполнена в научно-популярном стиле, чтобы донести до максимального круга специалистов в области цифровой обработки изображений фундаментальные вопросы человеческого зрения и видеообработки, происходящей в глазу, и обсудить наиболее массовые заблуждения в этой области.

Образ глаза как биологического аналога 130Mpix-камеры со специальными «ночными» режимами, сложившийся у наших современников, не выдерживает никакой критики. Огромное количество явных и неявных ошибок в научной и учебной литературе по поводу механизма зрения, и даже строения глаза, заставляет задуматься об объективных причинах этого явления. Например, сославшись на эпиграфы и проанализировав успехи наук о мозге, частью которого глаз безусловно является (а за последние 50 лет эти успехи более чем скромны [1,2,3,...] при колоссальных затратах!), придем к малоутешительному выводу о «встроенном» ограничении на самопознание. Но хотя бы поверхностный, инженерный, взгляд на проблему должен же быть нам позволен (!), чем и воспользуемся (в противном случае специалистам по видеообработке и быстродействующим процессорным системам было бы «и в зеркало глядя, трудно сохранить лицо»...).

Глаз и Видеопроцессор

Внесем, для начала, некоторую ясность для корректного понимания всем известной конструкции глаза, где двухлинзовый объектив из роговицы (+40 дптр) и перестраиваемого хрусталика (max +18 дптр) с диафрагмой радужной оболочки через прозрачную среду стекловидного тела и «водянистой влаги» (+4 дптр) проецирует перевернутое изображение на фоточувствительный слой, называемый сетчаткой (рис.1) [4,14]. Но вот далее – менее известные факты. Обратная (инверсная) сет-

«...И даже самый глаз не может, несмотря на совершенство строения, видеть самого себя!»

У.Шекспир

«Сетчатка – это наиболее передовая часть мозга.»

(перепраируя Дж.Даулинга)

чатка человека представляет собой 10-слойное сложное образование, где светочувствительные клетки (палочки и колбочки) находятся не впереди – торцом к свету, как многие полагают, а на самом дне, занимая менее 15 % толщины сетчатки и торцом упираются в темный (светопоглощающий) пигментный слой.

Светочувствительным слоем на дно – это даже в свое время стало предметом спора дарвинистов и креационистов: слишком неоптимальной представлялась подобная конструкция, как предмет творческого замысла. А модернизационный потенциал такой схемы обе стороны не могли осознать (споры продолжаются и поныне, причем высказывается масса претензий к разработчику(!), биолог-эволюционист Р. Дюкинс в 1986 году даже озаглавил свою книгу «Слепой часовщик»).

Таким образом, главная структура сетчатки (сюда следует включить также глазной нерв) – это Видеопроцессор Глаза, соединенный с одной стороны с фоторецепторами, а с другой – граничащий со стекловидным телом (далее Видеопроцессор). Сквозь его разводку и элементы коммутации (по иронии они называются «биполярыными клетками» – созвучно некоторым электронным схемам) свет, пройдя через хрусталик и стекловидное тело, теряясь и рассеиваясь, «пробирается» к боковым (!) поверхностям светочувствительных клеток, падает на них под острым углом и создает дифракционную картинку [5].

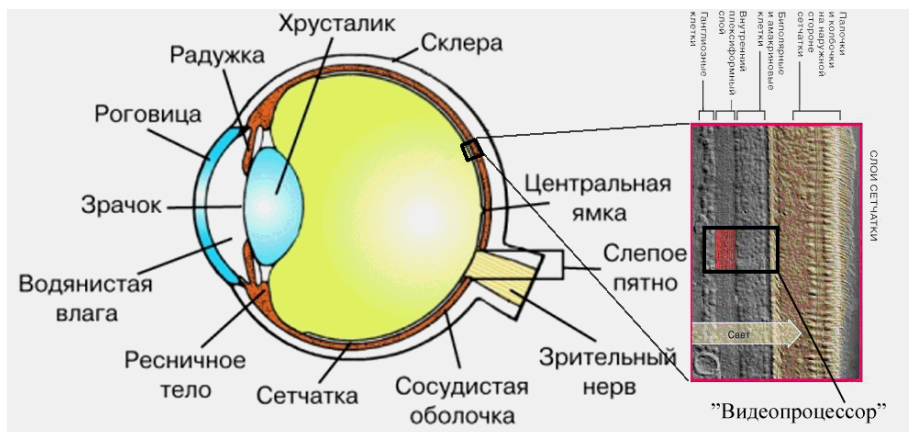


Рис. 1



Именно поэтому при анализе разрешающей способности глаза, так называемой «остроты зрения», неприемлем прямой «Рэлеевский» подход (угловое разрешение $1,22 \lambda/d$ – от длины волны и диаметра зрачка) – здесь уже массовая ошибка из учебников по физике и оптике, описывающих глаз: это не про глаз человека, разрешение которого много выше и, особенно, при сильно суженном зрачке!

Феномен зрения преподносит много элегантных решений – от конструкции фотосенсоров с выращиваемой в темноте (за 20 минут!) складчатой мембраной для получения чувствительности 10-20 фотонов и суммарного динамического диапазона более 10 порядков¹, до тонких решений кинематики и гидравлики. Но нас будет интересовать прежде всего видеобработка полученных от сенсоров сигналов возбуждения.

Обзор алгоритмов

Рассмотрим некоторые алгоритмы Видеопроцессора и попытаемся разделить их по функциональному признаку:

- Алгоритмы (программы) жизнеобеспечения (локальные контроллеры давления, температуры, влажности; управление механикой). Кроме того, многие ткани глаза, по-видимому, представляют собой биологические жидкие кристаллы (с потенциальным управлением?), и большинство известных неинфекционных болезней глаза, вероятно, можно свести к расстройствам, вызванным нарушениями в цепях обратной связи (контуры управления в области локальных контроллеров Видеопроцессора). Отсюда и феномены моментального излечения, когда внешнее воздействие (обычно ультразвук или ИК, но может быть и просто стресс!) восстанавливает обратную связь с участком, после чего сразу исчезают помутнения и другие симптомы. Это обнадеживает, хотя и не дает возможности вылечить многое «каплями».

- Калибровочные алгоритмы также жизненно важны для функционирования зрительной системы. Зрительный тракт во многих случаях является инструментом коррекции ошибок других сенсорных цепей (слух, тактильное восприятие и пр.), но до этого сам зрительный тракт должен быть хорошо откалиброван. Происходит это в первые годы жизни человека – только к полутора годам острота зрения ребенка достигает значения 0,5 (а в первые полгода – все видится как плохо собранный пазл, и весьма расплывчато). И дело здесь не столько в быстром росте и соответствующих геометрических проблемах глаза – аккомодации эластичного детского хрусталика было бы достаточно, чтобы с этим справиться, сколько – в сборке и настройке Видеопроцессора. При рождении заложена необходимая избыточность и структуры строятся как по генетическому проекту, формируя специализированные участки сетчатки (по функциональному виду клеток и типу обработки), так и под воздействием внешних стимулов, а главное – по результатам проведения необходимых калибровок по тест-объектам. Последнее особенно важно, чтобы разрешить совместную задачу настройки аккомодации и «фокусировки» дифракционной картинки. При этом требования к такому тест-объекту могут быть сформулированы следующим образом:

¹ $3 \cdot 10^{-4}$ лк – звезды без луны; $3 \cdot 10^3$ лк – операционное поле хирурга; $10^4 - 10^5$ лк – пляж на солнце.

- он должен быть достаточно крупным, несимметричным, характерной формы и содержать при этом мелкие детали для уточнения юстировки;

- образ его должен быть заранее занесен в память в нескольких проекциях;

- желательна априорная информация о его возможной угловой ориентации относительно оси глаза;

- объект должен располагаться в зоне оптимального зрения и расстояние до него должно быть известно (например, антропометрически) с хорошей точностью – чтобы разрешить противоречие при одновременной расфокусировке хрусталика и «сведении» дифракционной картинки Видеопроцессором.

И мы знаем такой уникальный объект (рис.2), притом с пропорциями «золотого сечения» – это кисть руки ребенка! (При этом основное положение при калибровке – рефлекторное мышечное, с открытой ладонью).



Рис.2

Недаром дети в раннем возрасте часами рассматривают свои ручки во всех проекциях, а изображение руки при этом проходит все стадии восприятия – от мутного подобия кленового листа (к ним многие потом неравнодушны!) до прецизионной «настроечной» таблицы с тонкими черточками перетяжек и четкими границами ногтей (хотелось бы подтвердить этот тезис авторов опытами, но психика ребенка слишком тонкая материя, чтобы вмешиваться). Потом, когда работа сделана, эта чудесная (местами, многократная!) избыточность в структуре сетчатки исчезает...

- Компенсирующие и адаптационные алгоритмы. Финалом калибровочных мероприятий являются компенсирующие алгоритмы, призванные всю дальнейшую жизнь прецизионно корректировать «конструкционные» недостатки глаза (еще на заре увлечения точной оптикой исследователь середины XIX века заметил, что, если бы ему принесли глаз как оптический прибор, он потребовал бы вернуть его изготовителю ввиду очень низкого качества оптики!).

К данному виду алгоритмов отнесем:

- широко известный «оборот изображения» (а на самом деле – цифровая компенсация индивидуальной, и местами хаотичной, схемы коммутации зрительного нерва, а не просто улучшение «двухлинзового объектива»);

- компенсация «слепого» пятна (место входа зрительного нерва);

- устранение менее известной «слепой» сетки (затенение «картинки» сосудами питания «процессорных слоев» в области кольцевого зрения);

- полная или частичная компенсация впоследствии приобретенных дефектов/травм (поражения сетчатки, кровоизлияния и пр.)

- компенсация насыщения засвеченных сенсоров (борьба с «блюмингом»).

Эффекты компенсации, по-видимому, достигаются посредством корреляции обработки с микроперемещением оптической оси (высокочастотная компонента так называемых микросаккад – частых произвольных (автоматических) [6] скачков глаза с частотой до десятков герц и амплитудой от нескольких угловых минут до 1°) и использованием памяти. Микросаккады при этом, помимо других, менее значимых с нашей точки зрения функций коррекции кинематики глаза, выполняют и более важную роль – «перебрасывают» одни и те же фрагменты изображения на разные, специализированные, участки нейронной сети (где могут в тот момент доминировать различные алгоритмы обработки), способствуя уточнению распознавания дифракционной картинки и всего фрагмента изображения. Этот механизм, наряду с тремором, вероятно и обуславливает существенное увеличение разрешающей способности зрения. Одновременно решается проблема компенсации влияния неоднородностей стекловидного тела и девиации чувствительности сетчатки. (Проблемы здесь аналогичны современным – в космических спутниках оптического зондирования тоже приходится компенсировать динамические неоднородности атмосферы. Результаты, достигнутые в Видеопроцессоре, могли бы иметь здесь хорошее применение!)

- Обеспечение безопасности (распознавание угроз, формирование команд, управление рефлекторными реакциями). Особенностью этих быстродействующих алгоритмов является допустимость большого процента ошибок – в соответствии с концепцией безопасности объекта. Много тысячелетий основной задачей такового было не попасть кому-нибудь в пасть и не встать на пути летящего предмета. К этому нужно добавить критерий защиты непосредственно глаза, для которого опасностью может быть и струя жидкости (яда), и температурная аномалия (ток воздуха) при пролете объекта (большого насекомого) вне поля зрения. Вот и сейчас, если показать человеку на экране большое контрастное пятно, которое вдруг резко увеличится в размере – он рефлекторно отпрянет, поскольку сработает один из таких алгоритмов (анализ производной угловой величины объекта). Но есть и более глубокое их влияние – поведенческое.

- Сложная видеобработка, предполагающая по обратной связи макроперемещение точки фокусировки при сканировании объектов (именно здесь видна траекторная разница при рассмотрении знакомых и новых объектов, а также реализация т.н. стратегии обзора – кстати, весьма индивидуальной для разных народов). Но это, главным образом, подготовка данных Видеопроцессором для зрительного отдела мозга – с дальнейшим использованием его большой памяти, распознаванием, обучением и пр. Все соотношения, перспективы и оценки расстояний, видимо, формирует также сам глаз [6], строго отмечая равенства расстояний и равенства от-

ношений. Например, перед любым прыжком (через яму!) делается автоматический анализ видимой перспективы и измерения неизвестных расстояний по отношению к известным. При движении же особенно «ценится» постоянство угловой величины видимости предмета. (Именно так утята, например, определяют дистанцию. Если размер утки впереди изменить в 1,5 раза – на столько же изменится дистанция, выбираемая утенком).

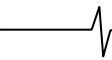
Отметим здесь, что попутно раскрывается и загадка притягательности «Золотого сечения» – это лишь реализация глобального «Принципа совпадения», всегда отмечаемого центром удовольствия. А здесь присутствует совпадение, и именно в равенстве отношений величин, экономная же легкость такого кодирования поощряется и воспринимается на неосознаваемом уровне как удовольствие (от оптимизации)!

- Наряду с несколькими, явно используемыми параллельно, типами обработки сигнала (что вообще характерно для обработки сенсоров мозгом человека – например, при обработке звука) следует обратить особое внимание на структуру глаза «два в одном». Имеется ввиду «центральное» и т.н. «кольцевое зрение», где по нашему мнению присутствуют разные типы обработки, в то время как периферическое зрение – просто специфично алгоритмами детектирования движения, обнаружения объектов с малой яркостью и пр. – это те же алгоритмы кольцевого зрения, но с повышением чувствительности в ущерб разрешению. (По нашему предположению количество сгруппированных сенсоров в сетчатке изменяется от центра поля к краю в приблизительном соответствии с «Золотым сечением» (рядом Фибоначчи) – на удалении от оси расположены кластеры из сотен объединенных палочек). В связи с этим периферическое зрение не выделено нами в отдельную систему.

Структура глазной ямки («фовеа») предполагает свой набор алгоритмов видеобработки для центральной области поля зрения (что подтверждает большое отличие топологии слоев процессора в этой зоне), а кольцевая зона вокруг фовеа – свой. Благодаря саккадам (название термина берет начало от «хлопка паруса» на старофранцузском (*saccade*) – то же, что и микросаккады, но с амплитудой более 1° , и меньшей частотой – до 2 Гц), достигается то, что переброс осуществляется уже в область кольцевого (и ночного) зрения (за пределы фовеа, обслуживающей лишь несколько градусов поля) с кардинально отличающимися алгоритмами, что особенно эффективно на границах и при смене освещенности. Таким образом, любой рассматриваемый объект при скачкообразном перемещении оптической оси неизбежно попадает под независимое (или детерминированное и заданное этим перемещением?) «изучение» обеими системами с применением двух и более типов обработки и всего арсенала алгоритмов – еще до того, как подключится зрительный отдел мозга.

Таким образом, в данной работе предлагается иерархическая модель зрения, с использованием микродвижений глаза на различных этапах обработки изображения, а именно:

- тремор – для «межпиксельной» обработки: уточнения центров дифракционных картинок точечных элементов изображения и обеспечения работы фотосенсоров без насыщения;



- микросаккады – для переброса элементов изображения между участками нейронной сети внутри и в окрестностях фовеа;

- саккады – для переноса элементов изображения из фовеа в область кольцевого и периферического зрения (и наоборот) для дополнительной обработки, в том числе в зонах «ночных» алгоритмов.

Таким образом, совокупность микродвижений глаза при рассмотрении объекта совсем не хаотична, а отражает «алгоритмическую» специфику топологии сетчатки индивида.

Далее в иерархии идут уже осознанные движения глаза, определяемые когнитивным процессом. При этом мозг содержит полную модель сетчатки, верифицируя поступающую информацию, управляя низкочастотными режимами кинематики глаза и назначая высокочастотные автоматические режимы – в соответствии с характеристикой объекта, текущими потребностями организации панорамного зрения и идущей работой над образами (со всей своей обширной памятью и творческим подходом к достраиванию изображений!). Как следствие, приходим к динамическому характеру синтетической картинке с обновлением фрагментов по запросу.

Мы не касаемся здесь задач стереоскопии и бинокулярности, где требуется согласованное управление осями зрения (и областями обработки?) с точностью в несколько угловых минут, а также присутствуют эффекты «конкуренции» полей зрения, диспаратности и пр. Это еще больше осложняет изучение такой двойной бинокулярной системы, на чем, ввиду краткости данной работы, придется рассмотрение закончить.

Рассуждения

Таким образом, имеем целый набор алгоритмов, требующих высокой производительности (малых времен обработки), что при ограниченной скорости передачи нервного возбуждения (до 100м/с) и инерционности исполнительской механики однозначно определяет необходимость их выполнения непосредственно в глазу – т.е. Видеопроцессором (с использованием соответствующей памяти для хранения примитивов и пр.). По данным последних исследований пропускная способность зрительного нерва составляет всего 10Мбод [13] и поток представляет собой хорошо обработанную и сжатую информацию. Здесь уместно процитировать Фрэнка Верблина [14]: « В целом наши выводы сводятся к тому, что специализированные нервные клетки, или нейроны, расположенные в толще сетчатки, передают нечто, что напоминает дюжину различных видеофильмов, транслируемых одновременно. Каждый такой видеопоток представляет собой непрерывное упрощенное отображение одного из аспектов наблюдаемой сцены, которую сетчатка непрерывно посылает в мозг. Например, один из зрительных потоков передает подобный карандашному наброску контур предмета, намечающий лишь границы объектов, другой реагирует на движение, причем часто лишь в одном определенном направлении, третьи каналы несут информацию о свете и тенях, а функции некоторых и вовсе трудно интерпретировать однозначно...». Вот так – параллельная обработка, задолго до появления первых процессорных комплектов, и при этом все те же противоречивые требования: при максимальном сжатии изображения, обеспечить резервирование линий передачи и избыточное кодирование![15]

Отсюда результаты «чтения мыслей», а точнее, зрительных образов, с затылочной области мозга электродами напрямую, без накопления картинки, получаются столь «крупнопанельными»: удается угадывать лишь главные черты композиции, а поток (RAW), видимо, рассыпан для надежности среди примитивов по тем самым 12-ти каналам, как это принято в нашей человеческой конструкции (см. например, кодирование в волокнах слухового нерва). Таким образом, большую роль играют кодированные образы – те самые примитивы, причем они архетипичны (по Юнгу) и специфичны для популяции, поскольку отражают, в том числе, ее историю. Понимание структуры примитивов позволило бы куда экономнее синтезировать изображение, оставляя его восприятие высококачественным (однако, ошибки здесь опасны и могут вызвать реальную головную боль и другие эффекты!).

Феномены и наблюдения

Практика – критерий истины, и должны быть не только сложные эксперименты, но и жизненные наблюдения, в той или иной мере подтверждающие высказанные здесь суждения. Сославшись в начале данной работы на ее популярный стиль, позволим себе в этом юбилейном издании использовать для ссылок несколько экзотические источники. Так, например, феномен невидимости ниндзя, помимо приемов суггестивной техники, находит здесь более очевидное объяснение. Для невидимости предмета достаточно, чтобы по нескольким основным потокам (например, оконтуривание, движение, яркость) он был отбракован из категории «объект» в категорию «шум». Но пороги при обработке, как и упоминавшиеся примитивы, национально ориентированы и такая техника может вовсе не сработать за пределами своей популяции (здесь сошлемся на известную любовь большинства древних правителей иметь иноземную охрану – быть может, не только ради осложнения процедуры подкупа?) и, памятуя юбилейный характер настоящего издания, приведем ниже документ о похождениях разведчиков-ниндзя времен русско-японской войны [16] (со ссылкой на музей г.Новочеркасск)². Становятся также понятнее манипуляции престижиджитаторов, жалобы водителей, «не увидевших» рядом другого автомобиля и многое другое.

Видеопроцессор и сетчатка

Вот почему в сетчатке глаза человека эта чрезвычайно загруженная задачами часть мозга выполнена с такой не свойственной другим органам расточительностью и при этом тщательностью и расчетом на будущие модернизации. Считается, что когда нет творчества, мозг работает на 5% (определяют по интенсивности кровоснабжения «территорий»). Но сетчатка – тот процент мозга, который на свету работает «на все 100%»: на единицу массы ткань сетчатки потребляет на 300% больше кислорода, чем ткани мозга и, как видим, не только ради процесса преобразования фотонов.

² От подьесаула Маркова Федора к Его Благородию полковнику Зимину А.Ф. О происшествии, случившемся в ночь на 17 августа 1904г.: «Мы сидели во второй линии охранения, жгли костер (линия фронта была достаточно далеко), варили еду. Вдруг из кустов выскочили два японца, обернутые во все черное, стали прыгать, всячески искривляться, как-то выть, шипеть и странно размахивать руками. Подошедшим к ним сотником Кривошлыковым наиболее шумный из басурман был ударен в ухо, отчего тут же и помер. Второй, убежав, вскоре был сыскан неподалеку же в кустах рядовым казаком Манилой Андреем, связан нагайкой и отнесен к палатке полкового есаула Вереева для снятия допроса».

Опишем биологический аспект Видеопроецессора хотя бы кратко: структура процессора в основной области сетчатки представляет собой трехслойную организацию с горизонтальными и вертикальными линейными и радиальными (в зависимости от типа клеток) связями многопортовых узлов. Очень многообещающая архитектура! Результаты одного из первых структурных исследований Видеопроецессора [7] (с датой на фотографии!) представлены на рис.3 [8,14] слева, а справа – увеличенный схематичный фрагмент (помещенный до этого в рамку на рис.1. и развернутый здесь).

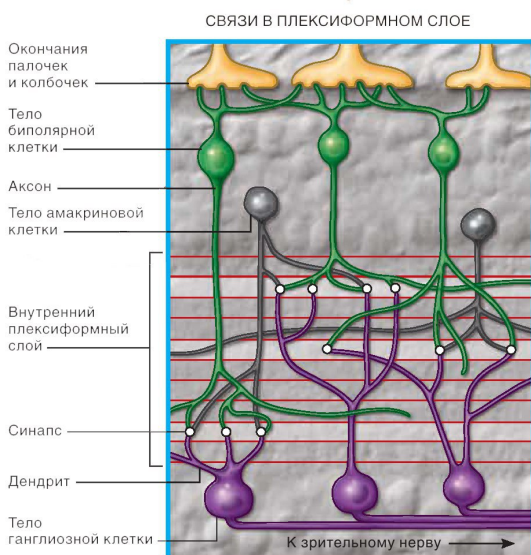
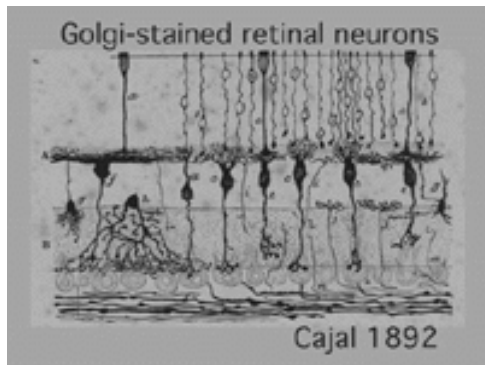


Рис.3

В свете представленных суждений вызывает большой интерес исторический аспект в части гистологии сетчатки – найденный замороженным в леднике глаз кроманьонца был бы бесценен для нас.

Космическая автономность

Отдельно следует остановиться на автономности такой системы как «государство Глаз», которая, с инженерной точки зрения, исключительна. Там присутствуют: система резервного внутреннего питания и регенерации расходных материалов; автономные системы защиты, аварийного отключения, поддержания давления, канализации, омывки и пр. – с минимальной зависимостью от внешней среды (при комфортной окружающей температуре). Потенциальная возможность для глаза существовать вне тела при подаче питания (хирургические

опыты со съемом фотозлектрической активности отдельно живущего глаза) это подтверждает. Ближайший современный аналог глаза по принципам и подходам к построению – космический аппарат при корабле-«матке». Следует предположить, что такая степень автономности обуславливает, вероятно, и самостоятельную, в том числе по временным параметрам, программу эволюции Видеопроецессора, как элемента Эволюции вида homo sapiens. Это, по-видимому, вполне согласуется с проектом так критикуемого разработчика – того самого, плохо разбирающегося в проблемах зрения и оптики Часовщика!...

Заключение

Воспользовавшись инженерным взглядом и категориями ЦОС, мы попробовали внести посильный вклад в работу медиков и физиологов (иногда трудную до безнадежности) по осмыслению некоторых фактов и, возможно, даже несколько опередили их в суждениях. Но стоит допустить, что основной конфуз при исследовании глаза (в отличие от других органов) состоит в том, что изучается изменяющийся, т.е. быстро обновляющийся объект. Результаты многих классических офтальмологических и психологических исследований середины 70-х сейчас могут уже не повториться, а данные современных обследований различных возрастных групп популяции, показывающие меньшую толщину слоя нейронов сетчатки у старших поколений, вполне вероятно могут ошибочно трактоваться, как возрастная дистрофия, влияние экологии и пр. – маскируя этим следы системного процесса.

Видеопроцессор, развернутый «внутри» глаза – резерв Эволюции. Незначительное увеличение толщины сетчатки за счет дополнительных нейронов и связей многократно увеличивает возможности обработки, лишь незначительно отражаясь на ухудшении ночного зрения, а возможно, компенсируется и это.

Следует предположить автономную эволюцию глаза – этой исключительной по уровню автономности системы организма, что по разным данным и наблюдаем последнее столетие. История наблюдений за остротой зрения и некоторые данные ретроспективы гистологии сетчатки позволяют сделать предположение о бурно идущей, но не очень заметной эволюции Видеопроецессора (отсюда – большой плюрализм в измерительных данных разных лет [4,7,9-12 и др.]). Например, различные источники дают рост числа колбочек и сопутствующих нейронов с 4-5 млн до 8-9 млн для европейцев в конце прошлого столетия: понятнее становятся конфликты зрительных предпочтений поколений и разная скорость работы «отцов и детей» за компьютером.

Направление Эволюции – развитие Видеопроецессора с модернизацией структуры и даже уменьшением, при необходимости, размера фотосенсоров. При этом толщина сетчатки и показатель остроты зрения в пределах одной популяции – косвенные критерии модернизации Видеопроецессора и индикаторы Эволюции. Она, видимо, уже началась (в XIX веке) и идет небывалыми для современной истории человечества темпами: философам и религиозным мыслителям здесь есть над чем задуматься!



Полемика приветствуется (на адрес редакции: e-mail info@dspa.ru).

Литература

1. Beaumont J.G. (ed.) Devided visual-field studies of cerebral organization. Academic, London, 1982
2. Красота и мозг. Биологические аспекты эстетики: пер. с англ. М., Мир,1995.
3. Бехтерева Н.П. Магия мозга и лабиринты жизни. М.,АСТ, 2007.
4. <http://www.krugosvet.ru/uploads/enc/images/t16/1236155023643b.jpg>
5. Хазен А.М. Разум природы и разум человека. М. НТЦ «Университетский», 2000.
6. Филин В.А. Автоматия саккад М.: МЦ «Видеоэкология». изд. Моск. Университет. 2001.
7. Cajal, Ramon y, Histologie du Systeme Nerveux de l'Homme et des Vertebres, trs. L. Azoulay, Paris: Maloine, 2 vols, 1909, 1911.
8. <http://webvision.med.utah.edu/OPL1.html>
9. Dowling, J. E., The Retina: An approachable part of the brain, Cambridge, MA:Harvard University Press, 1987.
10. Rodieck, R. W., 'The primate retina', Comparative Primate Biology, 4, 203-278, 1988.
11. Brockerhoff, S. E., Dowling J. E. and Hurley J. B, 'Zebrafish retinal mutants', Vision Research, 38, 1335-1339, 1998.
12. Hendrickson, A.E. and Youdelis, C. The morphological development of the human fovea. Ophthalmology 91, 603-612, 1984.
13. Kristin Koch, Judith McLean, Ronen Segev,Michael A. Freed,1 Michael J. Berry II,Vijay Balasubramanian, and Peter Sterling, How Much the Eye Tells the Brain. Current Biology 16, 1428–1434, July 25, 2006.
14. Frank Werblin, Botond Roska, The Movies in Our Eyes//Scientific American Magazine , №4, 2007.
15. Botond Roska, Alyosha Molnar, Frank S. Werblin, Parallel Processing in Retinal Ganglion Cells: How Integration of Space-Time Patterns of Excitation and Inhibition Form the Spiking Output. J Neurophysiol 95: 3810-3822, 2006.
16. Без традиций нет будущего (беседа с В.М.Бутровым)//Русская Традиция: Альманах. – М.,2001. – Вып.1,стр.148.

THE MAIN VIDEOPROCESSOR OF OUR CIVILIZATION – THE UNIQUE EXAMPLE OF THE EVOLUTION IN THE NEWEST HISTORY?

Kirpichnikov A.

The present article is written in popular scientific style and addressed to the wide range of experts in signal processing. It is devoted to the eye Videoprocessor and structure. Various signal processing algorithms which present at the eye and are specified by the functional properties and complex parallel processing and data preparation for the brain are mentioned. The necessity of the fulfilment these fast algorithms in the retina Videoprocessor directly is discussed.

The tremor, microsaccades and saccades as the hierarchical structure for images building are mentioned and the model of the retina in the brain as the source for the eye control commands is named. The location of the Videoprocessor ahead ("inverse retina") is the reserve of the Evolution. As the insignificant increase of the retina thickness does not particularly influence the optical parameters of the eye, but allows to multiple increase the ability of the signal processing.

The supposition concerning the autonomous Evolution of the eye is discussed. The direction of the Evolution is in increase the Videoprocessor power.

The indirect criteria of the eye Evolution in the modern Man History are named here as the increase of two factors – the retina thickness and visual acuity.

УЧЕБНАЯ ЛАБОРАТОРИЯ ЦИФРОВОЙ ОБРАБОТКИ СИГНАЛОВ РЯЗАНСКОГО ГОСУДАРСТВЕННОГО РАДИОТЕХНИЧЕСКОГО УНИВЕРСИТЕТА

Витязев С.В., н.с. кафедры ТОР РГРТУ, tor@rgta,ryazan.ru

Развивая рубрику журнала «ЦОС в университетах», мы начинаем обзор учебных лабораторий ЦОС, организованных и успешно действующих в ряде российских вузов. Обзор открывается рассказом о лаборатории, действующей при Рязанском государственном радиотехническом университете.

Отвечая целям и задачам широкого внедрения в учебный процесс цифровых информационных технологий реального времени, на базе кафедры Телекоммуникаций и основ радиотехники Рязанского государственного радиотехнического университета (www.rsreu.ru) в 2007 году была создана специализированная учебная лаборатория: «Цифровая обработка сигналов и DSP-технологии». Лаборатория располагает новейшим оборудованием, необходимым программным и учебно-методическим обеспечением по циклам дисциплин, связанным с применением цифровой обработки сигналов в радиотехнике, системах телекоммуникаций и управления, вычислительной технике, приборостроении, промышленной электронике. Учебная лаборатория дополнительно оснащена мультимедийным оборудованием, предоставляющим возможность чтения лекций с одновременным проведением практических и лабораторных занятий на 12 рабочих станциях.

Каждый учебный стенд включает в себя ПЭВМ, набор сменных учебных модулей DSP-разработчика ведущих зарубежных (Texas Instruments Inc. (США), Analog Devices Inc. (США)) и отечественных фирм (ГУП НПЦ «ЭЛВИС», ЗАО «Инструментальные системы», г. Москва), осциллограф, аудиоаппаратуру, необходимое лицензионное программное сопровождение. Лабораторные и практические занятия проводятся с широким применением среды MATLAB. Учебная лаборатория «Цифровая обработка сигналов и DSP-технологии» отвечает всем требованиям подготовки высококвалифицированных специалистов.

В лаборатории проводятся занятия по ряду дисциплин, связанных с цифровой обработкой сигналов и ее применением в различных отраслях науки и техники. Особо выделим два курса.

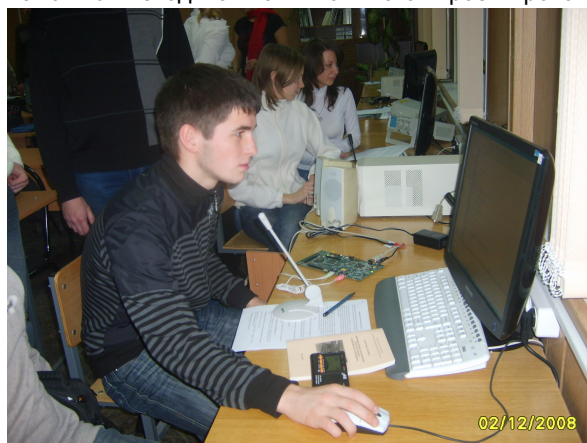
Цифровая обработка сигналов и сигнальные процессоры

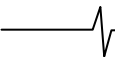
Курс рассчитан на два семестра, первый из которых является теоретическим и посвящен изучению основ теории цифровой обработки сигналов и оптимального проектирования систем ЦОС на сигнальных процессорах с применением среды MATLAB. Теоретическая часть курса включает в себя как классические разделы: дискретные сигналы и системы, дискретные преобразования, проектирование цифровых фильтров в классе КИХ- и БИХ-цепей, так и современные методы теории

обработки цифровых сигналов: многоскоростная и адаптивная обработка, вейвлет-преобразование, нелинейная фильтрация, обработка многомерных сигналов и изображений. Применение теории и методов ЦОС иллюстрируется многочисленными примерами. В зависимости от направления подготовки специалистов – это сжатие, кодирование, адаптивная эхо-компенсация и борьба с замираниями в системах телекоммуникаций; анализ и синтез сигналов, оптимальная фильтрация и пространственно-временная адаптивная обработка сигналов в радиотехнике; идентификация динамических объектов, цифровая фильтрация и оптимальный синтез управляющих сигналов в системах автоматического управления.



Изложение материала лекций и проведение лабораторных занятий сопровождается демонстрацией эффективности методов и алгоритмов ЦОС, реализуемых в среде MATLAB. Ряд разделов курса, и в первую очередь, - вопросы многоскоростной и адаптивной обработки сигналов базируются на новых, оригинальных результатах, полученных в процессе научных исследований, проводимых в лабораториях кафедры и университета. Отличительной особенностью курса является сочетание классических приемов и методов обработки сигналов с методикой оптимального проектирования





систем и устройств ЦОС на сигнальных процессорах заданного семейства.

Второй семестр ориентирован на изучение цифровых сигнальных процессоров (ЦСП) на примере процессоров семейства TMS320C67x фирмы Texas Instruments Inc (TI). Семейство TMS320C67x относится к классу процессоров с очень длинным командным словом (VLIW). В этом смысле данное семейство не относится к классическим по внутренней организации сигнальным процессорам, обычно используемым на начальном этапе обучения. Однако современные тенденции развития ЦСП требуют от разработчика знания мощных архитектур сегодняшнего дня, умения быстро переходить на новые платформы, навыков высокоуровневого программирования ЦСП. С этой точки зрения процессоры TMS320C67x вписываются в учебный процесс идеально. Они являются представителями наиболее передовой в вычислительном отношении платформы C6000 ЦСП с плавающей точкой фирмы TI. Их архитектура «дружественна» к языкам программирования высокого уровня. Кроме того, наборы разработчика DSKC6713 на базе данных процессоров, предоставляемые в рамках Университетской программы TI, являются очень демонстративными, обладая специализированными аппаратными узлами для ввода, вывода и обработки аудиосигналов. В процессе обучения студенты проходят путь от разработки структуры системы ЦОС в системе MATLAB и переноса ее непосредственно на аппаратную платформу DSKC6713 до отладки и оптимизации сгенерированного кода на низком уровне с использованием ассемблера TMS320C6713. Обучение строится на основе специально разработанного учебного пособия, которое доступно для скачивания на сайте www.dsps.ru/cosvuz. Работа студентов поощряется со стороны Университетской программы TI по программе ESPA.

Микропроцессорная техника управления в системах электрооборудования

Целью данного курса является знакомство студентов, проходящих обучение по специальности «Электрооборудование и электрохозяйство предприятий, организаций и учреждений (140610)», с наиболее интересными и практически значимыми идеями, получившими своё воплощение в изделиях ведущих производителей микропроцессорной техники. Курс разбит на две части и читается два семестра. Первая часть представляет собой непосредственное продолжение дисциплины «Основы вычислительной техники», в рамках которой студенты получают общее представление о микропроцессорах и ПЛИС. На протяжении первой части курса студенты знакомятся с особенностями архитектуры современных микроконтроллеров (МК) на примере сигнального МК TMS320F2808 и приобретают первые навыки в разработке программного обеспечения с использованием интегрированной среды «Code Composer Studio». Выбор МК TMS320F2808 компании «Texas Instruments» объясняется, с одной

стороны, тем, что МК этой серии отвечают современному уровню развития микропроцессорной техники, а с другой стороны, возможностью приобретения этих МК в рамках Университетской программы компании «Texas Instruments».

Вторая часть курса, в отличие от первой, носит в основном практический характер. Студенты выполняют более сложные лабораторные работы, в ходе которых изучают принципы оптимизации программ (циклическая адресация при построении цифровых фильтров, совмещение ассемблерных блоков с программированием на языке C и т.п.), учатся создавать проекты с использованием операционной системы реального времени «DSP/BIOS», а также знакомятся с особенностями визуального программирования в среде «MATLAB». Вторая часть курса завершается подготовкой курсового проекта. При выборе тем курсовых проектов учитываются пожелания студентов, поскольку зачастую курсовой проект получает своё дальнейшее развитие на этапе дипломного проектирования. Студенты разрабатывают программы для управления импульсными преобразователями напряжения, источ-



никами питания, электроприводом и т.д.

Учебная лаборатория «Цифровая обработка сигналов и DSP-технологии» при Рязанском государственном радиотехническом университете предлагает заинтересованным организациям сотрудничество по программе обучения специалистов основам теории цифровой обработки сигналов и практике программирования цифровых сигнальных процессоров фирм Texas Instruments, Analog Devices и сигнальных контроллеров «Мультикор» ГУП НПЦ «ЭЛВИС». Параллельно с непосредственной работой с конкретной аппаратной платформой могут быть прочитаны теоретические курсы по радиотехнике, радиолокации, обработке сигналов и изображений и другие. Предприятия могут быть заинтересованы в подготовке спецгрупп студентов с последующим их приемом на работу или в повышении квалификации своих сотрудников.

Научный руководитель лаборатории – д.т.н., профессор Витязев Владимир Викторович, заведующий кафедрой телекоммуникаций и основ радиотехники РГРТУ.

Контактная информация: тел./факс: (4912) 46-04-38, e-mail: tor@rgta.ryazan.ru