

ЦИФРОВАЯ ОБРАБОТКА СИГНАЛОВ

Научно-технический журнал
№ 4/2010

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. Мишенков,
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) 621-06-10, (495) 621-16-39
Факс: (495) 621-16-39
E-mail: nto.popov@mtu-net.ru
vityazev.v.v.@rsreu.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
Компьютерная верстка: А.Л. Витязева
Дизайн: Н.С. Осипова

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

Подписано в печать 14.01.11 г.
Формат 60x90/8.

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

Заказ № 7555. 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:

Алпатов Б.А., Стротое В.В.

Алгоритм оценки параметров геометрических
преобразований изображений при мультиспектральном
наблюдении 2

Бабаян П.В., Фельдман А.Б.

Обнаружение и оценка параметров воздушных объектов в
видеопоследовательностях на основе
кратномасштабной пространственной обработки 7

Алпатов Б.А., Блохин А.Н., Муравьев В.С.

Алгоритм обработки изображений для систем
автоматического сопровождения воздушных объектов.12

Бабаян П.В., Смирнов С.А.

Слежение за объектом на основе алгоритма
сопоставления с эталоном при одновременном
наблюдении в видимом и инфракрасном диапазонах 18

Бехтин Ю.С., Толстая Е.В.

Коррекция артефактов звона при JPEG-кодировании
изображений 22

Егошкин Н.А., Еремеев В.В.

Коррекция смаза изображений в системах космического
наблюдения земли 28

Побаруев В.И., Светелкин П.Н.

Хеш-тейловая пирамидальная реорганизация
изображений в системах обработки данных
дистанционного зондирования земли 33

**Антонушкина С.В., Еремеев В.В., Макаренко А.А.,
Москвитин А.Э.**

Особенности анализа и обработки информации от систем
гиперспектральной съемки земной поверхности 38

**Алпатов Б.А., Блохин А.Н., Костяшкин Л.Н.,
Романов Ю.Н., Шапка С.В.**

Семейство многофункциональных систем обработки
видеоизображений "Охотник" 44

**Вилкова Н.Н., Евстигнеев В.Г., Зубарев Ю.Б.,
Минаев В.Н.**

Новые технологии цифрового телевидения 52

Дворкович В.П., Федоров В.Ф.

Перспективы развития видеоконференцсвязи в России .. 57

Вилкова Н.Н., Быструшкин К.Н.

Российские инновационные технологии в приемной
аппаратуре цифрового телевидения на 22-ой
международной выставке "Связь-экспокомм-2010" 66

Зуев В.В., Жизняков А.Л.

Возможности использования фактора наследственности
признаков на последовательности изображений для их
анализа 69

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

АЛГОРИТМ ОЦЕНКИ ПАРАМЕТРОВ ГЕОМЕТРИЧЕСКИХ ПРЕОБРАЗОВАНИЙ ИЗОБРАЖЕНИЙ ПРИ МУЛЬТИСПЕКТРАЛЬНОМ НАБЛЮДЕНИИ

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

Ключевые слова: геометрические преобразования, мультиспектральное наблюдение.

Введение и постановка задачи.

Оценка параметров геометрических преобразований изображений является одной из важных задач, решаемых в системах видеослежения. Получаемые результаты используются при выделении движущихся объектов, оценке их координат, а также для электронной стабилизации изображений [1-3]. В современной литературе предложено достаточное количество алгоритмов решения данной задачи, в том числе ориентированных на работу в реальном масштабе времени [1,4].

Переход к мультиспектральному наблюдению изображений позволяет повысить надёжность функционирования систем видеослежения и расширить сферу их применения. При этом необходимо учитывать тот факт, что даже при неизменных во времени характеристиках светочувствительных матриц (линеек) и каналов передачи данных уровень шума для любого датчика может кардинально меняться во времени из-за изменения параметров наблюдаемой обстановки и работы системы АРУ. Следовательно, при поступлении изображений каждого кадра необходимо принимать решение, насколько целесообразно использование того или иного спектрального канала.

В данной работе предполагается, что все используемые датчики изображений съюстированы и формируемые изображения можно описать моделью:

$$I_{\gamma}(x, y, n) = g_{\gamma}(T(x, y, n)) + \xi_{\gamma}(x, y, n), \quad (1)$$

где $I_{\gamma}(x, y, n)$ – непрерывное изображение, наблюдаемое на кадре n в спектральном диапазоне γ ($\gamma = \overline{1, \Gamma}$), приведённое в процессе юстировки к системе координат изображения, наблюдаемого в первом из рассматриваемых спектральных диапазонов; $g_{\gamma}(x, y, n)$ – непрерывное изображение наблюдаемой обстановки; $\xi_{\gamma}(x, y, n)$ – аддитивный гауссовский белый шум.

Преобразование T описывает трансформацию, возникающую в процессе наблюдения сцены из удалённой точки. На первом кадре примем изображение g_{γ} нетрансформированным и ограничимся рассмотрением преобразования сдвига, однако заметим, что все приведённые рассуждения могут быть повторены применительно к преобразованиям других видов, например евклидовым, аффинным или проективным.

В работе рассмотрен многоэталонный алгоритм оценки параметров геометрических преобразований изображений, ориентированный на работу с мультиспектральными видеопоследовательностями. Предложен подход к выбору опорных участков изображений, позволяющий минимизировать ошибки, вызванные наличием аддитивных шумов на наблюдаемых изображениях. Представлены результаты экспериментальных исследований.

Алгоритм оценки параметров геометрических преобразований изображений при мультиспектральном наблюдении

При разработке алгоритма оценки параметров преобразования T будем опираться на известный из работы [5] алгоритм, работающий с одним спектральным каналом.

Принцип работы данного алгоритма состоит в следующем:

1. На исходном изображении выбираются опорные участки размером не более 32×32 пикселей;

2. При поступлении изображения очередного кадра на нём определяются положения всех опорных участков с субпиксельной точностью путём нахождения экстремума выбранной критериальной функции;

3. Из псевдорешения системы линейных уравнений, описывающей связь координат опорных участков в текущем и предыдущем кадрах, находятся предварительные оценки параметров межкадрового преобразования;

4. Так как на наблюдаемом изображении может находиться небольшое количество движущихся от кадра к кадру участков (например, содержащих изображения движущихся транспортных средств), которые могут быть выбраны в качестве опорных, возможно возникновение ошибок определения параметров преобразований. Для исключения таких ситуаций вычисляется вектор невязки упомянутой выше системы уравнений. На основе анализа элементов данного вектора из рассмотрения исключаются некоторые опорные участки и определяются уточнённые оценки параметров межкадрового преобразования;

5. На основе найденных, на предыдущем этапе оценок определяются параметры преобразования T ;

6. Производится замена опорных участков расположенных близко к границам изображения или имеющих большую невязку, после чего ожидается поступление изображения следующего кадра.

При переходе к мультиспектральному наблюдению возникают две задачи. Во-первых, необходимо решить, сколько опорных участков из общего их количе-

ства, определяемого аппаратными ресурсами системы, будет выбрано на изображении в каждом спектральном диапазоне в текущий момент времени и где конкретно они будут размещены. Во-вторых, необходимо корректно объединить результаты оценки параметров, полученные в различных спектральных диапазонах.

В том случае, когда первая задача решена, для решения второй задачи можно применять те же соотношения, что и в случае односпектрального наблюдения. Это возможно из-за того, что координаты всех опорных участков оказываются приведёнными к единой системе координат. Следовательно, необходимо рассмотреть задачу размещения опорных участков на изображениях $l_\gamma(x, y, n)$.

Выбор опорных участков при мультиспектральном наблюдении.

Выберем на изображении $l_\gamma(x, y, n_0)$ некоторый опорный участок. Его эталонное (незашумлённое) изображение обозначим $s(x, y, \gamma)$. Перечислим причины, которые могут приводить к ошибкам измерения его координат на текущем изображении $l_\gamma(x, y, n)$.

1. В области поиска опорного участка могут находиться объекты, изображения которых похожи на изображение опорного участка, вследствие чего искомые координаты будут определяться со значительной ошибкой.

2. Текущее изображение опорного участка искажено шумом, следовательно, положение экстремума критериальной функции определяется с некоторой погрешностью.

Ошибки, вызванные наличием похожих объектов, возникают достаточно редко. В рассмотренном выше алгоритме они могут быть обнаружены и парированы на этапе определения параметров межкадрового преобразования на основе анализ вектора невязки.

Напротив, ошибки, обусловленные наличием на изображении шума наблюдения, имеют меньшую величину (до нескольких пикселей), зависят от параметров эталонного изображения участка и возникают при каждом измерении координат. Следовательно, наилучшим опорным участком для использования в алгоритме оценки параметров геометрических преобразований будет участок, при оценке координат которого аддитивный шум будет вызывать наименьшую ошибку.

Необходимо отметить, что в реальных системах присутствуют ошибки, вызванные дискретной природой изображений, однако их величина мала (она составляет доли пикселя) и не зависит от параметров выбранного опорного участка.

Для того чтобы исключить ошибки, вызванные наличием похожих объектов при синтезе критерия выбора опорных участков, примем яркость всех точек изображения g_γ не принадлежащих участку $s(x, y, \gamma)$ равными нулю. Фактически, это означает переход от рассмотрения задачи определения положения опорного участка $s(x, y, \gamma)$ на изображении $l_\gamma(x, y, n)$ к рассмотрению задачи определения положения того же участка на изображении $z_s(x, y)$, определяемом выражением:

$$z_s(x, y) = s(x - \alpha_0, y - \beta_0, \gamma) + \xi_\gamma(x, y), \tag{2}$$

где (α_0, β_0) – координаты верхнего левого угла участка

S , $\xi_\gamma(x, y)$ – аддитивный шум в спектральном диапазоне γ . Нулевое значение полезного сигнала в точках, не принадлежащих участку S на изображении $z_s(x, y)$, позволяет оценивать координаты участка по положению максимума корреляционного интеграла [6]:

$$q(\alpha, \beta) = \int_0^W \int_0^H z_s(x, y) s(x - \alpha, y - \beta) dx dy = q_s(\alpha, \beta) + q_\xi(\alpha, \beta), \tag{3}$$

где из выражения (2)

$$q_s(\alpha, \beta) = \int_0^W \int_0^H s(x - \alpha_0, y - \beta_0) s(x - \alpha, y - \beta) dx dy,$$

$$q_\xi(\alpha, \beta) = \int_0^W \int_0^H \xi_\gamma(x, y) s(x - \alpha, y - \beta) dx dy, \tag{4}$$

а W, H – размеры изображений $l_\gamma(x, y, n)$ и $z_s(x, y)$. Разность измерения $(\tilde{\alpha}, \tilde{\beta})$ и истинного значения координат (α_0, β_0) представим в виде вектора

$$\Delta_{\alpha, \beta} = [\Delta\alpha, \Delta\beta]^T = [\alpha_0 - \tilde{\alpha}, \beta_0 - \tilde{\beta}]^T.$$

Для оценки точности вектор $\Delta_{\alpha, \beta}$ можно выразить через функции q_s , q_ξ и их частные производные при достаточно большом отношении сигнал/шум. Для этого предположим, что все частные производные изображения опорного участка до второго порядка включительно непрерывны, в том числе и на границе участка. Пользуясь данным предположением, выполним разложение в ряд Тейлора частных производных корреляционного интеграла $q(\alpha, \beta)$ в окрестности (α_0, β_0) , ограничившись рассмотрением первых двух членов. Упростим данное разложение, заметив, что в точке максимума функций $q(\alpha, \beta)$ и $q_s(\alpha, \beta)$ должно выполняться условие равенства нулю их первых частных производных. При большом отношении сигнал/шум значения вторых частных производных функции $q_\xi(\alpha, \beta)$ можно считать много меньшими значений остальных производных функций q_s и q_ξ , входящих в состав рассматриваемого разложения. Полученное в результате упрощения выражение может быть записано в виде:

$$H_{q_s}(\alpha_0, \beta_0) \cdot \Delta_{\alpha, \beta} = \nabla q_\xi(\alpha_0, \beta_0), \tag{5}$$

$$\text{где } H_{q_s}(\alpha_0, \beta_0) = \begin{bmatrix} \frac{\partial^2 q_s(\alpha, \beta)}{\partial \alpha^2} \Big|_{\substack{\alpha=\alpha_0, \\ \beta=\beta_0}} & \frac{\partial^2 q_s(\alpha, \beta)}{\partial \alpha \partial \beta} \Big|_{\substack{\alpha=\alpha_0, \\ \beta=\beta_0}} \\ \frac{\partial^2 q_s(\alpha, \beta)}{\partial \alpha \partial \beta} \Big|_{\substack{\alpha=\alpha_0, \\ \beta=\beta_0}} & \frac{\partial^2 q_s(\alpha, \beta)}{\partial \beta^2} \Big|_{\substack{\alpha=\alpha_0, \\ \beta=\beta_0}} \end{bmatrix},$$

$$\nabla q_\xi(\alpha_0, \beta_0) = \left[\frac{\partial q_\xi(\alpha, \beta)}{\partial \alpha} \Big|_{\substack{\alpha=\alpha_0, \\ \beta=\beta_0}}, \frac{\partial q_\xi(\alpha, \beta)}{\partial \beta} \Big|_{\substack{\alpha=\alpha_0, \\ \beta=\beta_0}} \right]^T.$$

Для оценки точности воспользуемся критерием среднего квадрата ошибки. Исходя из него, наилучшим будем считать тот опорный участок, для которого:

$$M \{ \Delta\alpha^2 + \Delta\beta^2 \} \Big|_{\alpha_0, \beta_0} \rightarrow \min. \tag{6}$$

Перепишем левую часть выражения (6), выразив вектор $\Delta_{\alpha, \beta}$ из выражения (5):

$$M \{ \Delta\alpha^2 + \Delta\beta^2 \} = M \{ \Delta_{\alpha, \beta}^T \cdot \Delta_{\alpha, \beta} \} = M \left\{ \nabla^T q_\xi(\alpha_0, \beta_0) \cdot (H_{q_s}^{-1}(\alpha_0, \beta_0))^T \cdot H_{q_s}^{-1}(\alpha_0, \beta_0) \cdot \nabla q_\xi(\alpha_0, \beta_0) \right\} \tag{7}$$

Рассмотрим подробнее элементы сомножителей в правой части выражения (7) под знаком математического ожидания. Для второй производной по α от информационной функции можно записать:

$$\begin{aligned} \left. \frac{\partial^2 q_s(\alpha, \beta)}{\partial \alpha^2} \right|_{\substack{\alpha=\alpha_0, \\ \beta=\beta_0}} &= \frac{\partial^2}{\partial \alpha^2} \left(\int_0^W \int_0^H s(x-\alpha_0, y-\beta_0) s(x-\alpha, y-\beta) dx dy \right) \Big|_{\substack{\alpha=\alpha_0, \\ \beta=\beta_0}} = \\ &= \int_0^H \int_0^W s(x-\alpha_0, y-\beta_0) \left(\frac{\partial^2 s(x-\alpha, y-\beta)}{\partial \alpha^2} \Big|_{\substack{\alpha=\alpha_0, \\ \beta=\beta_0}} \right) dx dy \end{aligned}$$

Принимая во внимание, что

$$\frac{\partial s(x-\alpha, y-\beta)}{\partial \alpha} = -\frac{\partial s(x-\alpha, y-\beta)}{\partial x},$$

$$\frac{\partial^2 s(x-\alpha, y-\beta)}{\partial \alpha^2} = \frac{\partial^2 s(x-\alpha, y-\beta)}{\partial x^2}, \quad (8)$$

Получаем

$$\left. \frac{\partial^2 q_s(\alpha, \beta)}{\partial \alpha^2} \right|_{\substack{\alpha=\alpha_0, \\ \beta=\beta_0}} = - \int_0^H \int_0^W \left(\frac{\partial s(x-\alpha, y-\beta)}{\partial x} \Big|_{\substack{\alpha=\alpha_0, \\ \beta=\beta_0}} \right)^2 dx dy. \quad (9)$$

Аналогично можно показать, что

$$\left. \frac{\partial^2 q_s(\alpha, \beta)}{\partial \alpha \partial \beta} \right|_{\substack{\alpha=\alpha_0, \\ \beta=\beta_0}} = 0; \quad (10)$$

$$\left. \frac{\partial^2 q_s(\alpha, \beta)}{\partial \beta^2} \right|_{\substack{\alpha=\alpha_0, \\ \beta=\beta_0}} = - \int_0^H \int_0^W \left(\frac{\partial s(x-\alpha, y-\beta)}{\partial y} \Big|_{\substack{\alpha=\alpha_0, \\ \beta=\beta_0}} \right)^2 dx dy. \quad (11)$$

Рассмотрим выражение для математического ожидания квадрата первого элемента вектора $\nabla q_\xi(\alpha_0, \beta_0)$:

$$\begin{aligned} M \left\{ \frac{\partial q_\xi(\alpha, \beta)}{\partial \alpha} \Big|_{\substack{\alpha=\alpha_0, \\ \beta=\beta_0}} \cdot \frac{\partial q_\xi(\alpha, \beta)}{\partial \alpha} \Big|_{\substack{\alpha=\alpha_0, \\ \beta=\beta_0}} \right\} = \\ \int_0^W \int_0^H \int_0^W \int_0^H M \{ \xi_\gamma(x_1, y_1) \xi_\gamma(x_2, y_2) \} \frac{\partial s(x_1 - \alpha, y_1 - \beta)}{\partial \alpha} \Big|_{\substack{\alpha=\alpha_0, \\ \beta=\beta_0}} \times \\ \times \frac{\partial s(x_2 - \alpha, y_2 - \beta)}{\partial \alpha} \Big|_{\substack{\alpha=\alpha_0, \\ \beta=\beta_0}} dx_1 dy_1 dx_2 dy_2. \end{aligned}$$

Учитывая, что для белого гауссовского шума

$$M \{ \xi_\gamma(x_1, y_1) \xi_\gamma(x_2, y_2) \} = \frac{\sigma_{\xi_\gamma}^2}{2} \delta(x_1 - x_2, y_1 - y_2),$$

где $\sigma_{\xi_\gamma}^2$ – дисперсия шума $\xi_\gamma(x, y)$, $\delta(x, y)$ – двумерная δ -функция, а также выражение (8), получаем

$$M \left\{ \left(\frac{\partial q_\xi(\alpha, \beta)}{\partial \alpha} \Big|_{\substack{\alpha=\alpha_0, \\ \beta=\beta_0}} \right)^2 \right\} = \frac{\sigma_{\xi_\gamma}^2}{2} \int_0^W \int_0^H \left(\frac{\partial s(x-\alpha, y-\beta)}{\partial x} \Big|_{\substack{\alpha=\alpha_0, \\ \beta=\beta_0}} \right)^2 dx dy. \quad (12)$$

Аналогично можно показать, что

$$M \left\{ \left(\frac{\partial q_\xi(\alpha, \beta)}{\partial \beta} \Big|_{\substack{\alpha=\alpha_0, \\ \beta=\beta_0}} \right)^2 \right\} = \frac{\sigma_{\xi_\gamma}^2}{2} \int_0^W \int_0^H \left(\frac{\partial s(x-\alpha, y-\beta)}{\partial y} \Big|_{\substack{\alpha=\alpha_0, \\ \beta=\beta_0}} \right)^2 dx dy. \quad (13)$$

Исходя из выражений (7)–(13) и учитывая, что элементы матрицы $H_{q_s}(\alpha_0, \beta_0)$ являются неслучайными, перепишем соотношение (6) в виде:

$$\begin{aligned} &\frac{\sigma_{\xi_\gamma}^2}{2 \int_0^W \int_0^H \left(\frac{\partial s(x-\alpha, y-\beta)}{\partial x} \Big|_{\substack{\alpha=\alpha_0, \\ \beta=\beta_0}} \right)^2 dx dy} + \\ &+ \frac{\sigma_{\xi_\gamma}^2}{2 \int_0^W \int_0^H \left(\frac{\partial s(x-\alpha, y-\beta)}{\partial y} \Big|_{\substack{\alpha=\alpha_0, \\ \beta=\beta_0}} \right)^2 dx dy} \rightarrow \min_{\alpha_0, \beta_0}. \end{aligned} \quad (14)$$

Полученное выражение показывает влияние параметров эталонного изображения участка $s(x, y)$ на величину ошибки.

Все рассуждения в данном разделе были проведены для изображений, представленных в непрерывном виде. Переходя к рассмотрению дискретных изображений можно утверждать, что при достаточно малом шаге дискретизации наиболее близким к наилучшему опорному участку будет участок, расположенный на изображении кадра в диапазоне $\hat{\gamma}$ в точке с координатами $\hat{\alpha}_0, \hat{\beta}_0$, определяемой следующим выражением:

$$\begin{aligned} (\hat{\alpha}_0, \hat{\beta}_0, \hat{\gamma}) = \\ = \arg \min_{\alpha_0, \beta_0, \gamma} \left(\frac{\sigma_{\xi_\gamma}^2}{\sum_{(i,j) \in S(\alpha_0, \beta_0, \gamma)} (\Delta_i s(i, j, \gamma))^2} + \frac{\sigma_{\xi_\gamma}^2}{\sum_{(i,j) \in S(\alpha_0, \beta_0, \gamma)} (\Delta_j s(i, j, \gamma))^2} \right), \end{aligned} \quad (15)$$

где $S(\alpha_0, \beta_0, \gamma)$ – множество точек участка S с координатами верхнего левого угла (α_0, β_0) на изображении $l_\gamma(i, j, n_0)$; запись вида $\Delta_i s$ означает операцию численного дифференцирования s по i ; $\sigma_{\xi_\gamma}^2$ – дисперсия аддитивного шума на изображении $l_\gamma(i, j, n_0)$.

Заметим, что выражение (15) не может быть использовано для выбора опорного участка в реальных системах из-за того, что эталонное изображение участка неизвестно, а для получения его незашумленной оценки необходимо наблюдать изображения данного участка на протяжении нескольких кадров. Поэтому, необходимо преобразовать выражение (15) таким образом, чтобы в нём фигурировало только наблюдаемое изображение опорного участка. При этом предполагается, что параметры шума не меняются во времени. Из модели (1) следует, что на множестве $S(\alpha_0, \beta_0, \gamma)$ верно следующее соотношение:

$$l_\gamma(i, j, n_0) = s(i, j, \gamma) + \xi_\gamma(i, j, n_0). \quad (16)$$

Выразим s из соотношения (16) и подставим в формулу (15). При этом знаменатель обеих дробей в правой части полученного выражения разделим на константу, равную количеству точек в области S равно $W \cdot H$. Упрощая данное выражение исходя из свойств аддитивного шума, получим окончательное выражение:

$$(\hat{\alpha}_0, \hat{\beta}_0, \hat{\gamma}) = \arg \min_{\alpha_0, \beta_0, \gamma} \left[\frac{\sigma_{\xi_y}^2}{\frac{1}{W \cdot H} \sum_{(i,j) \in S(\alpha_0, \beta_0, \gamma)} (\Delta_\gamma l_\gamma(i, j, n_0))^2 - 4\sigma_{\xi_y}^2} + \frac{\sigma_{\xi_y}^2}{\frac{1}{W \cdot H} \sum_{(i,j) \in S(\alpha_0, \beta_0, \gamma)} (\Delta_\gamma l_\gamma(i, j, n_0))^2 - 4\sigma_{\xi_y}^2} \right] \quad (17)$$

Достаточно точная оценка дисперсии шума может быть получена из сравнения изображений одного и того же участка, взятых из изображений соседних кадров:

$$\hat{\sigma}_{\xi_y}^2 = \frac{1}{2W \cdot H} \sum_{(i,j) \in S(\alpha_1, \beta_1, \gamma)} (l_\gamma(i, j, n_1) - l_\gamma(T_{n_1, n_2}^E(i, j, n_2)))^2, \quad (18)$$

где T_{n_1, n_2}^E – преобразование, связывающее кадры n_1 и n_2 .

Экспериментальные исследования

Были проведены сравнительные исследования двух алгоритмов: односпектрального алгоритма определения параметров преобразований, описанного в работе [5], работающего либо в телевизионном (TV), либо инфракрасном (IR) диапазоне и разработанного алгоритма, ориентированного на работу в обоих спектральных диапазонах одновременно (обозначен на рисунке 1 и в таблице 1 как TV+IR).

Исследования выполнены на натуральных мультиспектральных последовательностях изображений. Длительность видеопоследовательностей составляет 2 минуты при частоте поступления кадров равной 25 Гц, яркость точек изображений лежит в интервале от 0 до 1. Датчики, формировавшие изображения в обоих диапазонах были закреплены соосно на управляемой опорно-поворотной платформе, выполнена процедура юстировки изображений.

Общее количество опорных участков для каждого алгоритма равно 15, размер опорных участков равен 32×32 пикселя. Точность работы алгоритмов определяется ошибкой оценки равномерного субпиксельного сдвига (дрейфа), так величина добавленного при проведении исследования целочисленного сдвига определяется без ошибок.

На рисунке 1 приведены результаты исследований, которые показывают, что точность определения сдвига увеличивается при использовании обоих спектральных диапазонов. Это особенно хорошо заметно на графике 1(б). Графики 1(в) и 1(г) показывают, что хотя оцененный уровень шума на изображениях в телевизионном диапазоне (график 1(в), чёрная кривая) в основном выше, чем шум в тепловизионном диапазоне (график 1(в), серая кривая), тем не менее, наличие большего количества мелких деталей позволяет выбирать в телевизионном диапазоне больше опорных участков, нежели в тепловизионном (график 1(г)).

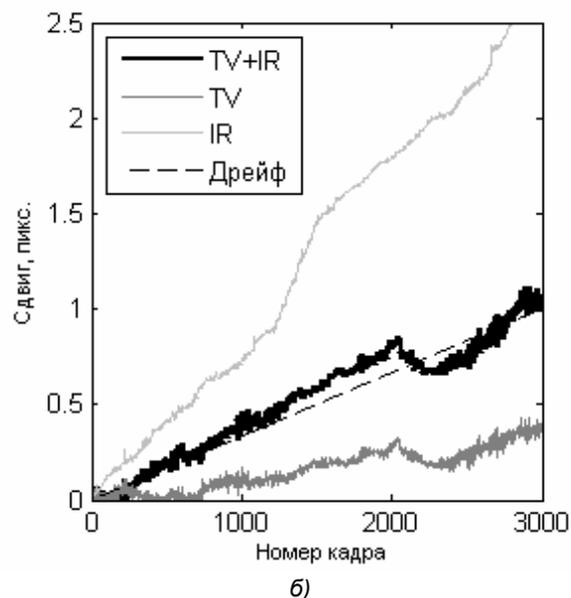
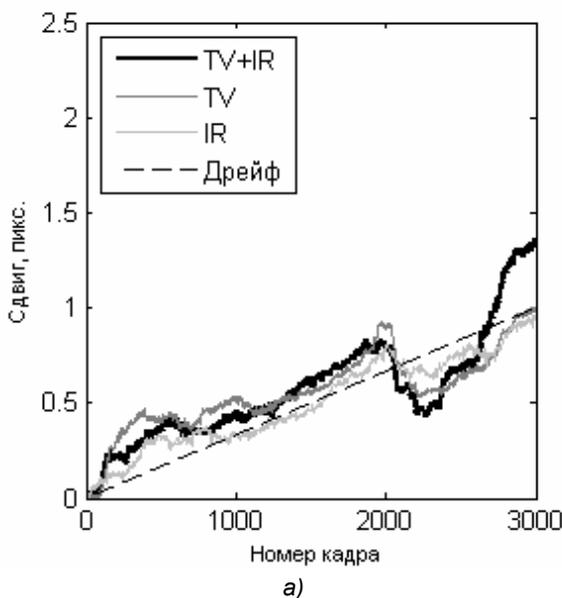
Были проведены исследования перечисленных выше алгоритмов при повышенном уровне шума. Для этого тестовые последовательности были зашумлены аддитивной помехой ζ_{add} различной интенсивности. Точность работы алгоритма оценивалась по усреднённому СКО ошибки определения координат сдвига. Результаты исследований, приведённые в таблице 1, показывают, что алгоритм, использующий оба спектральных диапазона, гораздо более устойчив к воздействию аддитивного шума.

Заключение

Разработанный многоэталонный алгоритм оценки параметров преобразований изображений используется для случая наблюдения изображений в нескольких спектральных диапазонах. Синтезирован критерий выбора опорных участков, учитывающий уровень шума и особенности изображений наблюдаемой обстановки в каждом из спектральных диапазонов.

Результаты проведённых исследований показывают, что системы обработки видеoinформации, оснащенные несколькими датчиками, работающими в различных спектральных диапазонах, обеспечивают большую точность при решении задач оценки параметров преобразований изображений, а также более устойчивы к воздействию аддитивного шума.

Исследования проводились при финансовой поддержке Министерства образования и науки РФ (госконтракт № 16.740.11.0223).



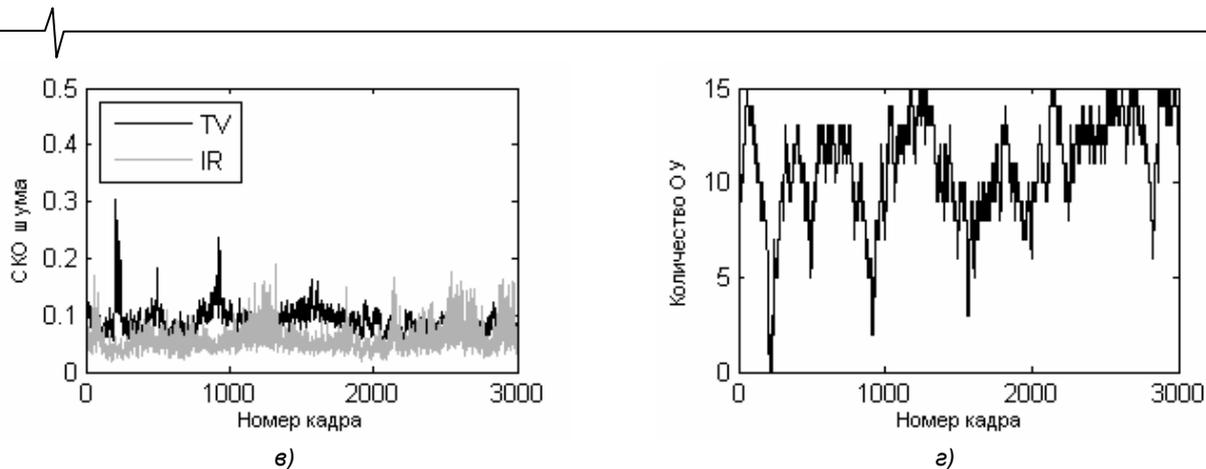


Рис. 1. Результаты исследований: оценки сдвига изображений по координатам X (а) и Y (б), оценки СКО шума на изображениях в различных диапазонах (в), количество участков, расположенных в ТВ диапазоне при работе мультиспектрального алгоритма (д).

Таблица 1

Оценка точности работы алгоритмов при различных уровнях аддитивного шума

$D_{\zeta_{add}}$	$3 \cdot 10^{-3}$	$4 \cdot 10^{-3}$	$5 \cdot 10^{-3}$	$6 \cdot 10^{-3}$	$7 \cdot 10^{-3}$	$8 \cdot 10^{-3}$	$9 \cdot 10^{-3}$
$\sigma_{ош} (TV+IR)$	0,24	0,24	0,26	0,30	0,38	0,50	0,59
$\sigma_{ош} (TV)$	0,44	0,63	1,20	1,32	1,97	2,49	3,20
$\sigma_{ош} (IR)$	2,44	3,15	3,02	3,28	3,67	3,95	4,46

Литература

1. Алпатов Б.А., Бабаян П.В., Балашов О.Е., Степашкин А.И. Методы автоматического обнаружения и сопровождения объектов. Обработка изображений и управление. – М.: Радиотехника, 2008. – 176 с.
2. Путятин Е.П., Аверин С.И. Обработка изображений в робототехнике. – М.: Машиностроение, 1990. – 320 с.
3. Zitova B., Flusser J. Image registration methods: a survey // Image and Vision Computing 21–2003, p.977–1000
4. Ташлинский А.Г. Оценивание параметров пространственных деформаций последовательностей изображений. – Ульяновск: УлГТУ, 2000. – 131 с.
5. Стротов В.В. Оценивание параметров смещения изображения в задачах выделения движущихся объектов // Вестник РГРТУ. Вып 23. – Рязань, 2008. – С. 30 – 37.
6. Баклицкий В.К., Бочкарёв А.М. Методы фильтрации сигналов в корреляционно-экстремальных системах

навигации. – М.: Радио и связь, 1986. – 216 с.

IMAGE GEOMETRIC TRANSFORMATION PARAMETERS ESTIMATION ALGORITHM FOR MULTISPECTRAL OBSERVATION

Alpatov B.A., Strotov V.V.

In this article an image geometric transformation parameters estimation algorithm which deals with multispectral videosequences is considered. An approach of choosing reference areas that allow to minimize caused by additive noise presence errors is proposed. The results of experimental examination are given.

Keywords: image processing, geometric transformations, multispectral observation

УДК 621.391.26

ОБНАРУЖЕНИЕ И ОЦЕНКА ПАРАМЕТРОВ ВОЗДУШНЫХ ОБЪЕКТОВ В ВИДЕОПОСЛЕДОВАТЕЛЬНОСТЯХ НА ОСНОВЕ КРАТНОМАСШТАБНОЙ ПРОСТРАНСТВЕННОЙ ОБРАБОТКИ

Бабаян П.В., к.т.н., начальник лаборатории Рязанского государственного радиотехнического университета, e-mail: aitu@rsreu.ru

Фельдман А.Б., аспирант Рязанского государственного радиотехнического университета, e-mail: aitu@rsreu.ru

Ключевые слова: видеопоследовательность, обнаружение объектов, измерение координат, оценка параметров, кратномасштабная обработка.

Введение

Одной из важных задач, решаемых в оптико-электронных системах наблюдения за воздушным пространством, является обнаружение и оценка параметров движущихся или неподвижных объектов, находящихся в поле зрения датчика изображений. При этом существует ряд особенностей, которые необходимо учитывать при решении данной задачи. Во-первых, датчик изображений может устанавливаться на поворотном устройстве и, следовательно, менять пространственную ориентацию. Во-вторых, фон, на котором наблюдаются объекты, представляет собой изображения ясного или облачного неба и содержит как сравнительно однородные участки, так и множество контрастных образований таких, как границы облаков, линия горизонта, деревья. В-третьих, часто отсутствует априорная информация о яркостных характеристиках и динамике наблюдаемых объектов.

На практике среди известных подходов к обнаружению воздушных объектов лучше всего зарекомендовали себя алгоритм статистической сегментации [1] и алгоритм, основанный на пространственной фильтрации изображений [2]. Среди недостатков алгоритма статистической сегментации, использующего различие выборочных распределений признаков для фона и объекта, можно выделить необходимость наличия априорных сведений о размерах и местоположении объекта и тот факт, что для малоразмерных объектов гистограмма не может служить хорошей оценкой распределения вероятности признаков. Вычислительная сложность алгоритма сегментации возрастает с увеличением числа используемых признаков, особенно при больших порядностях гистограмм. Основным недостатком алгоритма пространственной фильтрации, в котором для обнаружения объектов выполняются операции свертки изображения с масками специального вида с последующей пороговой обработкой и анализом полученных сегментов, можно считать тот факт, что при увеличении размера объекта увеличиваются и размеры апертур используемых фильтров.

Рассмотрен алгоритм обнаружения и оценки параметров воздушных объектов, особенностью которого является использование нескольких масштабных каналов с пониженной частотой пространственной дискретизации в каждом. Изображения в отдельных каналах рассматриваются как реализации марковского случайного поля. Для адаптации к изменениям наблюдаемых размеров объекта в каждом кадре осуществляется выбор наиболее подходящего масштабного канала. Представлены результаты сравнительных исследований предложенного подхода и алгоритма, основанного на пространственной фильтрации изображений. Проведен анализ вычислительной сложности для указанных алгоритмов.

Таким образом, в настоящее время актуальной является задача разработки алгоритма обнаружения малоразмерных воздушных объектов, обладающего низкими вычислительными затратами. При этом целесообразно учитывать наличие статистической связи между близлежащими точками фонового изображения, поэтому в настоящей работе предлагается рассматривать фоновое изображение как реализацию марковского случайного поля. На основании марковских моделей может быть синтезирован оптимальный алгоритм обнаружения точечных объектов. Для получения изображения, на котором интересующий объект представлен одним элементом разложения, используется процедура понижения частоты пространственной дискретизации, при этом коэффициент децимации меняется в процессе наблюдения для адаптации к изменениям наблюдаемого размера объекта.

Алгоритм обнаружения точечных объектов. Предположим, что фоновое изображение $g(i, j)$ является реализацией марковского случайного поля вида [3]

$$g(i, j) = \frac{1}{A} \sum_{(u,v) \in s(i,j)} \alpha_{(i,j)}^{(u,v)} g(u, v) + \theta(i, j), \quad A = \sum_{(u,v) \in s(i,j)} \alpha_{(i,j)}^{(u,v)}, \quad (1)$$

где $i = \overline{1, N_x}$, $j = \overline{1, N_y}$ – пространственные координаты; N_x, N_y – соответственно, число строк и столбцов изображения; $s(i, j) = \{(u, v) | 0 < (i - u)^2 + (j - v)^2 \leq 2\}$ – определяет множество элементов восьмисвязной окрестности для элемента (i, j) ; коэффициенты $\alpha_{(i,j)}^{(u,v)}$ определяют степень влияния пикселя (u, v) на пиксель (i, j) ; $\theta(i, j)$ – некоррелированный по пространственным координатам гауссовский случайный процесс с нулевым средним и дисперсией σ_0^2 (так называемый формирующий шум).

Пусть коэффициенты $\alpha_{(i,j)}^{(u,v)}$ зависят только от взаимного расположения пикселей (u, v) , (i, j) и не зависят от выбора точки (i, j) . Вопросы оценивания коэффициентов $\alpha_{(i,j)}^{(u,v)}$ обсуждаются, например, в статье [3].

Будем считать, что объект непрозрачный, и яркость объекта распределена равномерно $p(h) = 1/l_{\max}$, $h \in [0, l_{\max}]$, где l_{\max} – максимальное значение яркости изображения. Кроме того, положим, что дисперсия аддитивного шума датчика пренебрежимо мала по сравнению с дисперсией формирующего шума модели σ_{θ}^2 . В таком случае можно показать [4], что оптимальное по критерию Неймана-Пирсона правило принятия решения о присутствии объекта в точке наблюдаемого изображения $l(i, j)$ может быть записано в виде

$$\left| l(i, j) - \frac{1}{A} \sum_{(u,v) \in S(i,j)} \alpha_{(i,j)}^{(u,v)} l(u, v) \right| \begin{cases} \geq \hat{\sigma}_{\theta} \cdot k(p_-) \Rightarrow \text{объект;} \\ < \hat{\sigma}_{\theta} \cdot k(p_-) \Rightarrow \text{фон.} \end{cases} \quad (2)$$

Здесь $k(p_-)$ – квантиль уровня $p_-/2$ стандартного нормального распределения, p_- – заданная вероятность ложных тревог; $\hat{\sigma}_{\theta}$ – максимально правдоподобная оценка СКО формирующего шума

$$\hat{\sigma}_{\theta} = \sqrt{\frac{1}{N_x N_y} \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} \left(l(i, j) - \frac{1}{A} \sum_{(u,v) \in S(i,j)} \alpha_{(i,j)}^{(u,v)} l(u, v) \right)^2}. \quad (3)$$

Результатом применения решающего правила (2) в текущем кадре является множество точек, часть которых соответствует ложным тревогам, поэтому выполняется дополнительный анализ с целью устранить нестабильные во времени точки.

Оценка параметров точечного объекта. Будем считать, что на рассматриваемом промежутке времени датчик изображений либо неподвижен, либо характер его движения известен априори. В первом случае оценивание параметров объекта можно вести непосредственно в системе координат наблюдаемого изображения. Во втором случае целесообразно перейти в систему координат, связанную с неподвижным фоном. Для простоты изложения ограничимся рассмотрением ситуации, когда датчик изображений неподвижен.

При оценивании параметров заданного объекта в текущем кадре $\mathbf{L}_n = \{l(i, j, n)\}$ можно сократить зону поиска (строб) $\mathbf{S}_n \subset \mathbf{L}_n$ с учетом ограничений на максимальную скорость объекта. При отсутствии других точечных объектов в стробе следует ожидать, что наибольшее отличие между заданной точкой и ее соседями обусловлено присутствием интересующего объекта. Введем критериальную функцию

$$F(i, j, n) = \left| l(i, j, n) - \frac{1}{A} \sum_{(u,v) \in S(i,j)} \alpha_{(i,j)}^{(u,v)} l(u, v, n) \right|. \quad (4)$$

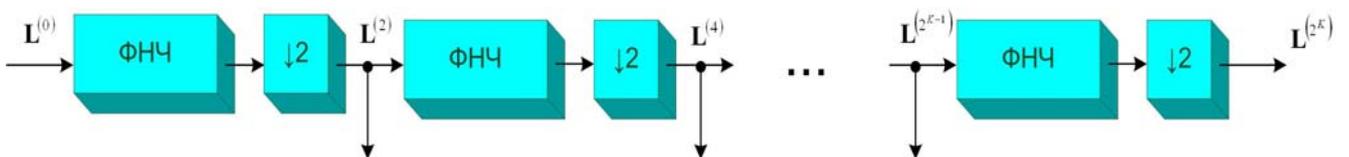


Рис. 1. Иллюстрация многоступенчатой процедуры децимации изображений (ФНЧ – фильтр нижних частот; $\mathbf{L}^{(2^k)}(i, j)$, $k = 0, K$ – прореженные изображения)

Тогда измерение координат $\tilde{\lambda}(n) = [\tilde{\lambda}_x(n), \tilde{\lambda}_y(n)]$ в текущем кадре может быть найдено по правилу

$$\tilde{\lambda}(n) = \arg \max_{(i,j) \in \mathbf{S}_n} F(i, j, n). \quad (5)$$

В силу эффектов пространственной дискретизации, чересстрочной развертки, наличия неучтенного аддитивного шума датчика и ряда других причин измерение координат по правилу (5) может сопровождаться ошибкой. Поэтому для уточнения измерений координат в текущем кадре, оценивания проекций вектора скорости объекта и позиционирования центра строба (центра области \mathbf{S}_{n+1}) в следующем кадре используются методы калмановской фильтрации, позволяющей принимать решение о соответствии измерений координат, полученных по правилу (5), интересующему объекту. Для этого результаты измерений в текущем кадре сравниваются с прогнозируемыми значениями координат, и, если отличие между указанными величинами превышает некоторый порог, зависящий от дисперсии оценок координат, принимается решение о пропуске объекта. В таком случае в качестве оценок параметров объекта используются прогнозируемые значения, формируемые фильтром Калмана. Если на протяжении заданного числа кадров не получено измерений координат, соответствующих объекту, принимается решение о потере объекта, другими словами – фиксируется срыв слежения.

Адаптация к размерам объекта путем децимации изображений. Если объект не является точечным и известны его начальные линейные размеры, то можно свести задачу к случаю наблюдения точечного объекта путем понижения частоты пространственной дискретизации наблюдаемого изображения. Экспериментальные исследования показали, что целесообразно выбирать коэффициенты прореживания кратные степеням числа 2. Для уменьшения вычислительных затрат при больших значениях коэффициента децимации понижение частоты дискретизации можно организовать ступенчато в несколько этапов, на каждом из которых изображение прореживается в два раза по каждой координате. Иллюстрация многоступенчатой процедуры децимации представлена на рис. 1.

Если результат оценивания координат объекта на прореженном в v раз по строкам и столбцам изображении $\hat{\lambda}_x^{(v)}, \hat{\lambda}_y^{(v)}$, то соответствующие оценки в системе координат исходного изображения – $\hat{\lambda}_x = v\hat{\lambda}_x^{(v)}, \hat{\lambda}_y = v\hat{\lambda}_y^{(v)}$.

Так как размеры объекта меняются в процессе наблюдения, может возникнуть ситуация, когда при данном коэффициенте децимации объект на прореженном изображении перестанет быть точечным.

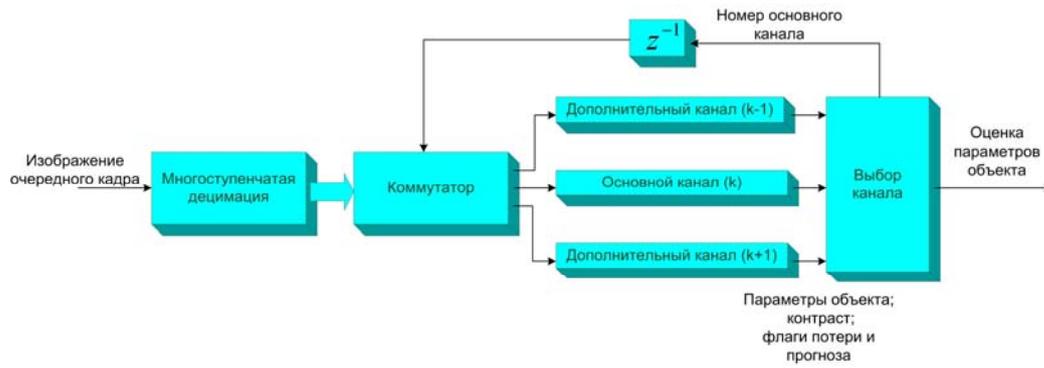


Рис. 2. Структурная схема многоканального алгоритма оценки параметров воздушных объектов

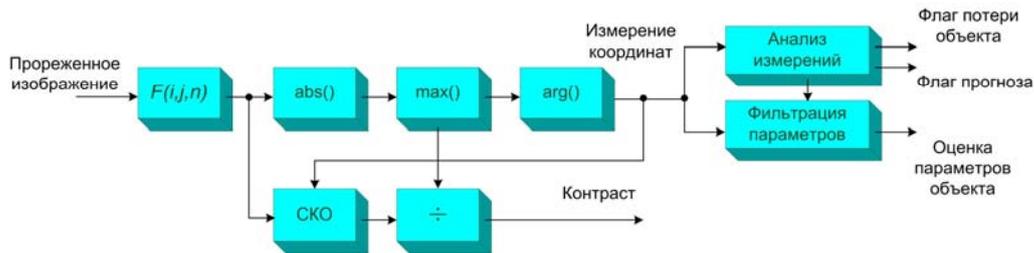


Рис. 3. Структурная схема масштабного канала

Учитывая, что размеры объекта не могут измениться скачкообразно, целесообразно вести обработку в трех масштабных каналах с близкими значениями коэффициента децимации (рис. 2).

На первом кадре в качестве основного выбирается канал k с коэффициентом децимации $v_k = 2^k$, ближайшим к линейным размерам объекта. Одновременно с основным включаются два соседних канала с коэффициентами прореживания v_{k-1} и v_{k+1} . Каналы работают независимо друг от друга, и на каждом кадре осуществляется выбор канала, наилучшим образом соответствующего наблюдаемым размерам объекта.

На рис. 3 представлена структура одного масштабного канала. Обработка информации осуществляется следующим образом. Прежде всего, выполняется вычисление критериальной функции (4), что можно эффективно реализовать в виде линейной свертки изображения с маской, элементы которой зависят от весовых коэффициентов модели (1). Далее производится измерение координат объекта согласно правилу (5). Измерения поступают на блок анализа, который при необходимости «выставляет» флаги потери и прогноза, и на блок фильтрации параметров, осуществляющий уточнение

координат и скорости объекта. Кроме того, в каждом канале вычисляется локальный контраст – отношение величины максимума отсчетов к СКО значений критериальной функции.

На каждом кадре необходимо принимать решение о том, какой из каналов наилучшим образом соответствует размерам объекта. Данную задачу решает блок выбора канала. Рассмотрим процедуру переключения на примере двух каналов – основного (ОК) и одного дополнительного (ДК). Соответствующая структурная схема представлена на рис. 4.

Критерием, указывающим на необходимость смены каналов, является превышение контрастом в дополнительном канале контраста, вычисленного в основном канале. При этом следует учесть, что максимум в дополнительном канале может не соответствовать интересующему объекту, поэтому применяется дополнительный анализ, который заключается в сопоставлении нескольких последних отсчетов траекторий и скоростей объектов в основном и дополнительном каналах. Кроме того, для принятия решения о переключении необходимо, чтобы дополнительный канал не находился в режиме потери или прогноза.

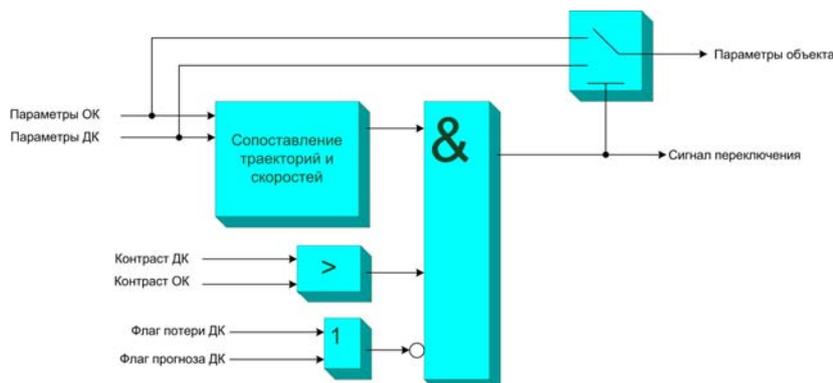


Рис. 4. Структурная схема блока выбора канала (на примере двух каналов: ОК – основной канал; ДК – дополнительный канал)



Пример работы алгоритма слежения на основе кратномасштабного подхода представлен на рис. 5. Обратим внимание на момент, когда наблюдаемый объект пересекает плотную облачность (рис. 5-г). Здесь энергия светового поля от объекта рассеивается облаком, что приводит к увеличению наблюдаемых размеров объекта. Реагируя на этот факт, алгоритм переключается на канал с коэффициентом децимации 4. Затем, когда объект выходит за пределы плотной облачности (рис. 5-д), происходит возврат к каналу с коэффициентом децимации равным 2.

Сравнительные экспериментальные исследования алгоритмов слежения за объектами, основанных на процедуре оценки параметров. Для исследования эффективности предложенного подхода к оцениванию параметров воздушных объектов (кратномасштабный алгоритм А1) были проведены сравнительные исследования на натуральных видеопоследовательностях. В качестве альтернативы использовался алгоритм оценивания параметров объектов, основанный на пространственной фильтрации изображений [2] (алгоритм пространственной фильтрации А2). В ходе исследований рассчитывались следующие количественные характеристики: 1) частота правильного обнаружения; 2) частота срывов слежения; 3) среднеквадратическая ошибка оценивания координат.

Под *частотой правильного обнаружения* понимается отношение числа кадров, на которых объект был об-

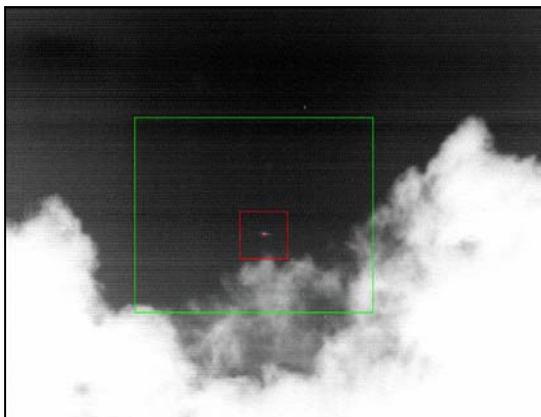
наружен, к общему числу кадров видеопоследовательности. При этом считается, что объект обнаружен, если разность между эталонными и измеренными координатами центра объекта не превышает диагонали прямоугольника, ограничивающего эталонное изображение объекта в данном кадре видеопоследовательности.

Частотой срывов слежения считается отношение числа ситуаций, когда в данном кадре фиксируется потеря объекта, к общему числу кадров видеопоследовательности.

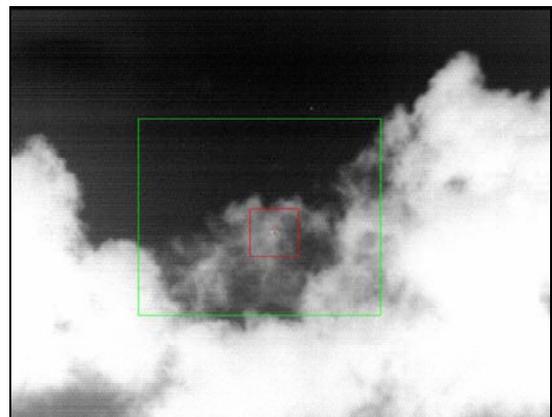
Под *среднеквадратической ошибкой оценивания координат* понимается модуль вектора, компонентами которого являются среднеквадратические ошибки оценивания соответствующих координат объекта.

Для проведения экспериментальных исследований была сформирована *коллекция тестовых видеосюжетов*, состоящая из 16 видеосюжетов, полученных в видимом и инфракрасном диапазонах спектра. Размеры наблюдаемых объектов колебались от 2x2 до 30x30 пикселей.

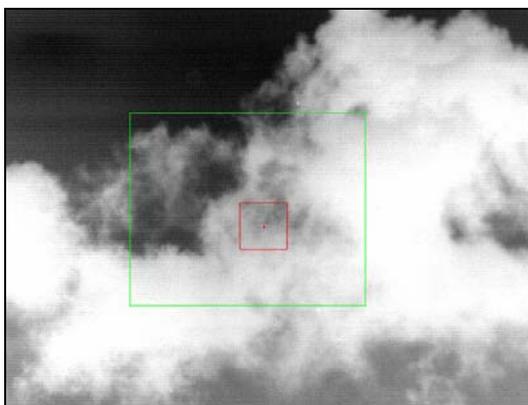
Рассмотрение *результатов экспериментов*, представленных в табл. 1, позволяет утверждать, что описанный в настоящей работе алгоритм А1 по частоте правильного обнаружения приближается к алгоритму А2. Между тем, как и ожидалось, алгоритм А1 демонстрирует меньшую точность, что объясняется работой с прореженными изображениями и отсутствием этапа выделения точек объекта.



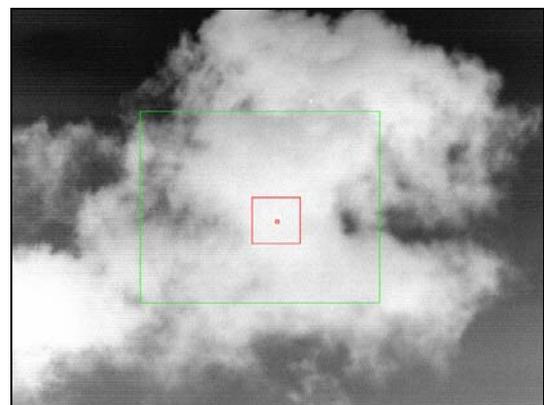
а) Кадр 5. Начало слежения за объектом



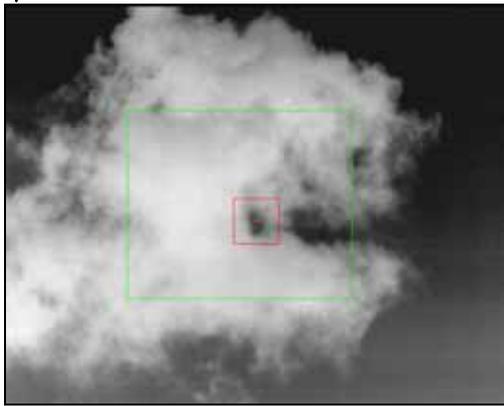
б) Кадр 100. Слежение за объектом на фоне перистой облачности



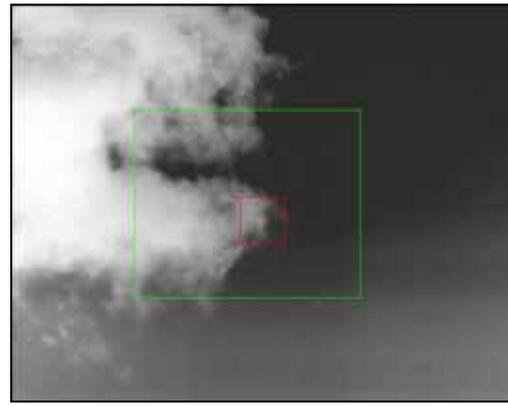
в) Кадр 200. Слежение за объектом на фоне облака



г) Кадр 300. Объект влетает в плотную облачность. Переход к каналу с коэффициентом децимации 4



д) Кадр 350. Возврат к каналу с коэффициентом децимации 2



е) Кадр 450. Слежение за объектом вблизи границы облачности

Рис. 5. Пример работы алгоритма слежения на основе кратномасштабного подхода к оцениванию параметров объекта

Таблица 1

Результаты сравнительного исследования алгоритмов слежения за воздушными объектами

Алгоритм	Частота правильного обнаружения	Частота срывов слежения	Модуль среднеквадратической ошибки, пикселей
A1	0,94	$1,3 \cdot 10^{-2}$	2,41
A2	0,97	$5,9 \cdot 10^{-3}$	1,22

Таблица 2

Число операций умножения с накоплением для исследуемых алгоритмов

Алгоритм	Линейные размеры объекта (в пикселях)				
	2	4	8	16	32
A1	$14,4 \cdot N$	$15,1 \cdot N$	$13,8 \cdot N$	$13,4 \cdot N$	$13,3 \cdot N$
A2	$20 \cdot N$	$20 \cdot N$	$20 \cdot N$	$36 \cdot N$	$104 \cdot N$

Анализируя вычислительную сложность исследуемых алгоритмов, учтем, что основным этапом каждого из них является линейная свертка. Число операций умножения с накоплением в зависимости от линейных размеров объекта при заданном размере изображения $N = N_x \cdot N_y$ для каждого из алгоритмов приводится в табл. 2. Отметим, что алгоритм A2 включает также этапы сегментации отфильтрованного изображения, разметки и параметризации полученных сегментов [5]. Вычислительная сложность этих этапов не зависит от размера объекта и поэтому не учитывалась при составлении табл. 2. Таким образом, из представленных данных видно, что алгоритм A1 имеет меньшую вычислительную сложность, которая еще более уменьшается с увеличением размера объекта, в то время как для алгоритма A2 объем вычислений, наоборот, возрастает. Алгоритм A1 обеспечивает выигрыш также по объему требуемой памяти, так как необходимо хранить только прореженные в несколько раз изображения для каждого из трех активных каналов.

Работа проводилась при финансовой поддержке Министерства образования и науки РФ (госконтракт № 02.740.11.0002).

Литература

1. Бакут П.А., Лабунец В.Г. Телевизионная следящая система с байесовским дискриминатором цели // Зарубежная радиоэлектроника. – 1987 – №10. – С.81-93.
2. Муравьев В.С. Пространственный алгоритм обнаружения и определения координат воздушных объектов на изображении

// Вестник Рязанского государственного радиотехнического университета. – Рязань, 2009. – №2(28). – С.17-20.

3. Perez P. Markov Random Fields and Images // CWI. – 1998. – №4 – PP.413-437.
4. Алпатов Б.А., Бабаян П.В., Фельдман А.Б. Обнаружение и оценка параметров малоразмерных воздушных объектов в последовательности изображений // Цифровая обработка сигналов и ее применение: Труды Российского научно-технического общества радиотехники, электроники и связи им. А.С.Попова. – М., 2010. – С.117-120.
5. Алпатов Б.А., Бабаян П.В., Балашов О.Е., Степашкин А.И. Методы автоматического обнаружения и сопровождения объектов. Обработка изображений и управление. – М.: Радиотехника, 2008. – 176 с.: ил.

AIR OBJECT DETECTION AND PARAMETER ESTIMATION IN THE IMAGE SEQUENCES ON THE BASIS OF MULTISCALE SPACE PROCESSING

Babayan P.V., Feldman A.B.

In this paper the air object detection and parameters estimation algorithm is proposed. The main feature of it is the use of the several scale channels with the reduced spatial frequencies in each. In the separated channel the observed image is considered as the sample of the Markov random field. One is able to select the channel with the best correspondence to the size of the object.

АЛГОРИТМ ОБРАБОТКИ ИЗОБРАЖЕНИЙ ДЛЯ СИСТЕМ АВТОМАТИЧЕСКОГО СОПРОВОЖДЕНИЯ ВОЗДУШНЫХ ОБЪЕКТОВ

Алпатов Б.А., зав. каф. АИТУ, д.т.н., профессор, Рязанский государственный радиотехнический университет, aitu@rsreu.ru

Блохин А.Н., нач. лаборатории, Государственный Рязанский приборный завод, hunter-rzn@yandex.ru

Муравьев В.С., научный сотрудник, Рязанский государственный радиотехнический университет, aitu@rsreu.ru

Введение

Постоянное совершенствование датчиков изображений и средств вычислительной техники приводит к расширению сферы практического применения методов и алгоритмов обработки изображений. Так, комплексы обработки изображений реального времени начали активно применяться в системах автоматического обнаружения и сопровождения воздушных объектов. Область использования технологий автоматического выделения и оценки параметров воздушных объектов включает в себя задачи обеспечения безопасности аэропортов и важных промышленных предприятий, контроль и мониторинг воздушного движения, астронавигацию и т.д.

Типовая структура системы автоматического сопровождения воздушных объектов приведена на рис. 1.

В системах автоматического сопровождения объектов датчик изображений обычно размещается на поворотном устройстве, прикреплённом к корпусу носителя, что позволяет изменять направление оптической оси датчика. Смещения поля зрения датчика могут приводить к существенным сдвигам изображения. Формируемые последовательности изображений также могут быть искажены шумами, источниками которых являются датчик, устройство квантования и др. Начальным этапом обработки видеоинформации, как правило, является этап выделения объектов, состоящий в получении би-

Рассматривается алгоритм выделения воздушных объектов на телевизионных и тепловизионных изображениях, ориентированный на реализацию в системах автоматического сопровождения реального времени. Представлена типовая структура системы автоматического сопровождения объектов, введена модель фоноцелевой обстановки, синтезирован алгоритм выделения объектов, предложен модифицированный алгоритм, опирающийся на результаты теоретических исследований, приведены результаты экспериментальных исследований.

нарной маски, определяющей положение и форму объектов на наблюдаемом изображении. Предполагается, что площадь наблюдаемых объектов на изображении может лежать в диапазоне от нескольких пикселей до нескольких тысяч пикселей. На практике в процессе выделения объектов любой алгоритм будет допускать ошибки, которые возможно уменьшить путем применения процедур логической или морфологической фильтрации бинарных изображений [1, 2]. После измерения параметров связанных областей полученной бинарной маски принимается решение об обнаружении интересных объектов. Потребителю выдается список найденных объектов и их параметров, к числу которых можно отнести яркость, размеры, координаты центра и др. Устройство управления необходимо для минимизации рассогласования между реальным направлением оптической оси датчика и желаемым, которое в большинстве случаев совпадает с направлением на объект [1].

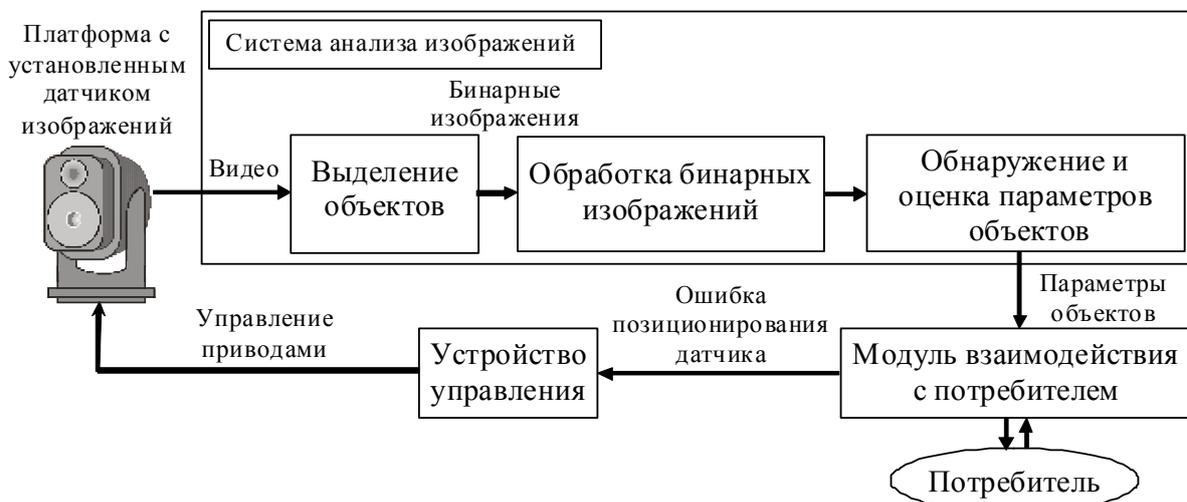


Рис. 1. Структура системы автоматического сопровождения воздушных объектов

Очевидно, что надежность сопровождения во многом зависит от качества бинарной маски объекта, полученной на этапе выделения. Следует упомянуть ряд подходов к решению задачи выделения объектов. Так, многие методы основываются на использовании пространственно-временной фильтрации [3, 4]. При этом из-за движения датчика изображения возникает необходимость в оценке параметров геометрических преобразований. В общем случае из-за неизвестных перемещений датчика изображений сложно установить соответствие между точками соседних кадров, поэтому применение алгоритмов, использующих временную фильтрацию, часто бывает затруднительным. Разработаны оптимальные алгоритмы выделения объектов, предполагающие известной корреляционную матрицу фона [5]. Однако при больших размерах изображения требуемая операция нахождения обратной матрицы является очень ресурсоемкой. В [6] предложен алгоритм выделения объектов, наблюдаемых на облачном небе, на основе нелинейной пространственной фильтрации. Существенным ограничением является то, что данный алгоритм способен выделять только малоразмерные объекты.

Таким образом, в силу ограничений существующих подходов разработка вычислительно эффективных алгоритмов выделения воздушных объектов, способных надежно работать в случае смещения поля зрения датчика и изменения размеров объектов, до настоящего времени остается актуальной задачей.

Постановка задачи

Рассмотрим аналитическую постановку задач выделения, обнаружения и измерения параметров объектов. Пусть монохроматический датчик формирует последовательность изображений с частотой $1/\tau$ кадров в секунду, где τ – период следования кадров. На кадре с номером n наблюдаемое изображение будет описываться яркостной функцией $l(i, j, n)$. Здесь и далее i, j – координаты пикселя, $(i, j) \in R$, $R = \{(i, j) : i = \overline{0, N_x - 1}; j = \overline{0, N_y - 1}\}$; R – множество точек, на котором задано изображение; N_x, N_y – размеры изображения по вертикали и горизонтали. Предполагается, что как объект, так и фон являются полностью “непрозрачными” для диапазона частот, в котором принимается оптический сигнал, атмосфера считается полностью прозрачной. Также необходимо учитывать шум, вносимый датчиком изображения, который, как правило, имеет аддитивный характер. С учётом сделанных замечаний модель наблюдения можно представить в виде [7]:

$$l(i, j, n) = h(i, j, n)r(i, j, n) + g(i, j, n)(1 - r(i, j, n)) + \xi(i, j, n), \quad (1)$$

$$(i, j) \in R, n = \overline{1, N},$$

где N – количество наблюдаемых кадров, $l(i, j, n)$ – наблюдаемое изображение, $g(i, j, n)$ и $h(i, j, n)$ – неизвестные яркости точек фона и объектов соответственно на кадре с номером n , $\xi(i, j, n)$ – шум с нулевым средним, некоррелированный по пространству и времени, функция $r(i, j, n)$ определяет расположение объектов на изображении и задаётся правилом:

$$r(i, j, n) = \begin{cases} 1, & \text{если в точке } (i, j) \text{ кадра } n \text{ находится объект,} \\ 0, & \text{в противном случае.} \end{cases} \quad (2)$$

Задача *выделения* объекта состоит в оценивании бинарной маски $r(i, j, n)$ по наблюдаемому изображению $l(i, j, n)$. Задача *обнаружения* объекта состоит в принятии решения о присутствии объекта на изображении. При этом $r(i, j, n)$ изменяется с течением времени вследствие изменения ракурса, расстояния до объекта, а также в связи с заслонениями объектов. На основе бинарной маски $r(i, j, n)$ можно вычислить такие параметры объекта, как левая, правая, верхняя и нижняя границы. Эти границы определяют прямоугольник, описанный вокруг объекта. Координатами центра объекта в кадре n будем считать координаты центра прямоугольника $(\alpha_{об}, \beta_{об})$, ограничивающего объект, а размерами объекта – высоту и ширину прямоугольника $(V_{об}, W_{об})$.

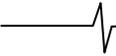
Для успешного решения задач выделения и обнаружения объектов необходимо располагать как можно более полной информацией о функциях яркости объектов и фона, а также знать параметры распределения шума. Оптимальный в смысле выбранного критерия качества алгоритм можно построить только при наличии полной априорной информации о свойствах объекта и фона. В большинстве случаев требуемая информация бывает недоступна. Реальные условия всегда характеризуются той или иной степенью неопределенности в отношении информации о фоновой обстановке, которая обусловлена как невозможностью точного предсказания свойств объекта и фона, так и непредвиденными изменениями этих свойств во времени. В ряде случаев можно задаваться некоторыми ограничениями на вид неизвестных функций $g(i, j, n)$ и $h(i, j, n)$.

Необходимо принимать во внимание, что из-за возможного движения датчика изображений формируемые кадры видеопоследовательности подвергаются воздействию геометрических преобразований. Известные алгоритмы оценки параметров геометрических преобразований [3, 8] требуют высоких вычислительных затрат и не всегда бывают эффективны при наблюдении облачных фонов, характеризующихся наличием протяженных коррелированных яркостных образований. Возможным выходом из этой ситуации является применение алгоритмов, использующих преимущественно пространственные модели изображения фона и объекта. Поэтому для упрощения записи будем опускать аргумент n при дальнейшем изложении.

При наблюдении удаленного объекта на равномерном фоне яркости объекта и фона в кадре можно считать неизвестными случайными величинами с плотностями распределения

$$p(h) = p(g) = \frac{1}{l_{\max}},$$

при $h \in [0; l_{\max}]$, $g \in [0; l_{\max}]$. Месторасположение объекта на изображении неизвестно, но предположим, что доступна информация о его конфигурации. Координаты (α, β) , задающие местоположение объекта на изображении, будем считать дискретными случайными величинами, подчиняющимися равномерному закону распределения. Количество возможных положений объекта в кадре равно $N_x \alpha N_y \beta$, где $N_x \alpha = N_x - V_{об} + 1$, $N_y \beta = N_y - W_{об} + 1$, $(V_{об}, W_{об})$ – соответственно высота и ширина ограничивающего



объект прямоугольника. Будем считать, что $\xi(i, j, n)$ – белый гауссовый шум с нулевым средним и известной дисперсией σ_ξ^2 .

Решение задачи

Для выбора между гипотезой X_1 о присутствии объекта и альтернативной гипотезой X_0 составляется отношение правдоподобия Λ , которое сравнивается с порогом. С учетом введенной модели фоноцелевой обстановки (1) принимает вид:

$$X_1: l(i, j) = \begin{cases} h + \xi(i, j), & (i, j) \in H_{об}, \\ g + \xi(i, j), & (i, j) \in G_{об}; \end{cases} \quad (3)$$

$$X_0: l(i, j) = g + \xi(i, j), \quad (i, j) \in R,$$

где $H_{об}$ – множество, задающее координаты точек объекта на изображении, $G_{об} = R \setminus H_{об}$. Выражение для Λ будет равно [9]:

$$\Lambda = \frac{\sum_{\alpha=0}^{N_{xx}-1} \sum_{\beta=0}^{N_{yy}-1} p(\alpha, \beta) \int_0^{l_{\max}} \int_0^{l_{\max}} p_1(L | g, h, \alpha, \beta) p(h) p(g) dh dg}{\int_0^{l_{\max}} p_0(L | g) p(g) dg}, \quad (4)$$

где $p_1(L | g, h, \alpha, \beta)$ – многомерная плотность распределения наблюдаемого изображения при наличии объекта, $p_0(L | g)$ – многомерная плотность распределения наблюдаемого изображения при отсутствии объекта, L – вектор, состоящий из яркостей точек наблюдаемого изображения.

С учетом сделанных допущений, оптимальное по критерию Неймана-Пирсона правило принятия решения будет заключаться в проверке следующего условия [9]:

$$\frac{1}{N_{xx} N_{yy}} \sum_{\alpha=0}^{N_{xx}-1} \sum_{\beta=0}^{N_{yy}-1} \exp \left(\frac{1}{2\sigma_\xi^2} \frac{S_g S_h}{N_x N_y} (\hat{h}_{\alpha, \beta} - \hat{g}_{\alpha, \beta})^2 \right) > C/k_\Lambda, \quad (5)$$

где C – константа, зависящая от заданной величины ложной тревоги,

$$k_\Lambda = \frac{\sqrt{2\pi}\sigma_\xi \sqrt{N_x N_y}}{l_{\max} \sqrt{S_h S_g}},$$

S_h – площадь объекта, $S_g = N_x N_y - S_h$, $\hat{h}_{\alpha, \beta}$, $\hat{g}_{\alpha, \beta}$ – усредненные значения яркостей объекта и фона, вычисляемые для всех α и β в областях, определяемых конфигурациями точек объекта и фона.

Из (5) видно, что для каждого положения объекта необходимо вычислять экспоненту, что при больших размерах изображения требует больших вычислительных затрат. Для уменьшения вычислительной сложности алгоритма, принимая во внимание неравенство $e^x > 1 + x$, перейдем к следующему правилу:

$$\sum_{\alpha=0}^{N_{xx}-1} \sum_{\beta=0}^{N_{yy}-1} (\hat{h}_{\alpha, \beta} - \hat{g}_{\alpha, \beta})^2 > \frac{2N_{xx} N_{yy} N_x N_y}{S_g S_h} (C/k_\Lambda - 1) \sigma_\xi^2, \quad (6)$$

что не приведет к возрастанию вероятности ложных тревог. Количественные значения вероятностей правильного обнаружения P_{no} и ложных тревог P_{lm} можно найти из следующих выражений:

$$P_{no} = 1 - \Phi \left(\frac{\sqrt{2N_{xx} N_{yy}} (C_1 - b/2)}{\sqrt{(1 + 2b)}} \right), \quad (7)$$

$$P_{lm} = 1 - \Phi \left(\sqrt{2N_{xx} N_{yy}} C_1 \right),$$

где $\Phi(\dots)$ – интеграл вероятностей,

$$C_1 = C/k_\Lambda - 3/2, \quad b = \frac{S_g S_h}{N_x N_y} SNR^2,$$

SNR – отношение сигнал/шум, вычисляемое как отношение контраста объекта к среднеквадратическому отклонению шума. Графики зависимостей P_{no} от P_{lm} , определяемых в параметрическом виде выражениями (7), при разных значениях SNR приведены на рис. 2.

Помимо принятия решения о присутствии искомого объекта на изображении необходимо находить оценки $(\hat{\alpha}_{об}, \hat{\beta}_{об})$ координат, определяющих его истинное местоположение. В соответствии с критерием максимального правдоподобия алгоритм решения данной задачи сводится к максимизации критериальной функции

$$F(\alpha, \beta) = \left| \hat{h}_{\alpha, \beta} - \hat{g}_{\alpha, \beta} \right|$$

и сравнении максимального значения с порогом, зависящим от площади объекта, размеров изображения и СКО шума. В случае принятия гипотезы о наличии объекта в кадре его координаты на изображении будут определяться как

$$(\hat{\alpha}_{об}, \hat{\beta}_{об}) = \arg \max_{\alpha, \beta} (F(\alpha, \beta)) \quad [9].$$

В реальных условиях наблюдения в кадре могут присутствовать несколько объектов, конфигурация которых неизвестна. Поскольку изображения могут быть искажены воздействием шума различной природы, а фон может быть как ровным, так и содержать разнородную облачность, то получить аналитическое решение задачи выделения объектов в общем случае не удастся. Поэтому предлагается внести ряд модификаций в рассмотренный выше алгоритм.

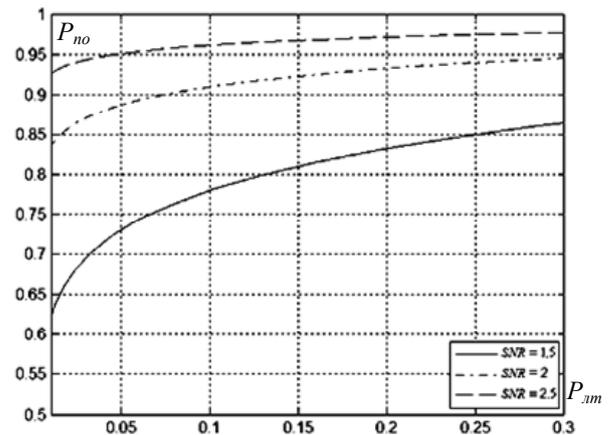


Рис. 2. Графики теоретических зависимостей P_{no} от P_{lm} при разных значениях SNR

Если $g(i, j)$ является изменяющимся процессом, то целесообразно ограничивать область, задаваемую $G_{об}$, некоторой прямоугольной зоной размерами $(x_g \times y_g)$, в пределах которой $g(i, j)$ можно считать по-

стоянной. Размеры зоны зависят от яркостной изменчивости фоновой составляющей и площади наблюдаемого объекта. В частности размеры прямоугольной зоны не должны быть меньше размеров объекта на изображении, т.е. должны выполняться условия $x_g > V_{об}$ и $y_g > W_{об}$. В большинстве случаев значение σ_ξ^2 также заранее неизвестно и для принятия решения о наличии объекта необходимо вычислять оценку $\hat{\sigma}_\xi^2$. С учетом вышеизложенного предлагаемый алгоритм выделения объектов заключается в следующем.

Входное изображение $l(i, j)$ обрабатывается линейным фильтром, имеющим маску w_1 размерности $q_{12} \times q_{11}$. Параллельно с этим $l(i, j)$ сглаживается фильтром с маской w_2 размерностью $q_{22} \times q_{21}$, причем $q_{21} > q_{11}, q_{22} > q_{12}$:

$$f_1(i, j) = \sum_{m_x=-(q_{12}-1)/2}^{(q_{12}-1)/2} \sum_{m_y=-(q_{11}-1)/2}^{(q_{11}-1)/2} w_1(m_x, m_y) l(i-m_x, j-m_y),$$

$$f_2(i, j) = \sum_{m_x=-(q_{22}-1)/2}^{(q_{22}-1)/2} \sum_{m_y=-(q_{21}-1)/2}^{(q_{21}-1)/2} w_2(m_x, m_y) l(i-m_x, j-m_y),$$
(8)

где $f_1(i, j)$ и $f_2(i, j)$ – выходные изображения, полученные после фильтрации масками w_1 и w_2 соответственно. При этом $f_2(i, j)$ представляет собой оценку фоновой составляющей наблюдаемого изображения. Маски фильтров имеют следующий вид:

$$w_1(m_x, m_y) = 1/q_{11}q_{12}, m_x = \overline{-(q_{12}-1)/2, (q_{12}-1)/2}, m_y = \overline{-(q_{11}-1)/2, (q_{11}-1)/2},$$

$$w_2(m_x, m_y) = \begin{cases} 0, & \text{при } m_x = \overline{-(q_{12}-1)/2, (q_{12}-1)/2}, m_y = \overline{-(q_{11}-1)/2, (q_{11}-1)/2}, \\ 1/(q_{21}q_{22} - q_{11}q_{12}), & \text{иначе.} \end{cases}$$
(9)

Во многих практических случаях маски фильтров можно выбирать квадратными, т.е. $q_{11} = q_{12} = q_1$ и $q_{21} = q_{22} = q_2$. Далее путем вычитания отфильтрованных изображений находится разность $d(i, j) = f_1(i, j) - f_2(i, j)$, что имеет много общего с вычислением выражения, стоящего в левой части (6). Итоговое решающее правило имеет вид:

$$\hat{r}(i, j) = \begin{cases} 1, & |d(i, j)| > k\hat{\sigma}, \\ 0, & \text{иначе,} \end{cases}$$
(10)

где k – пороговый коэффициент,

$$\hat{\sigma} = \sqrt{\frac{1}{N_x N_y} \sum_{i=0}^{N_x-1} \sum_{j=0}^{N_y-1} d^2(i, j)}$$

– оценка среднеквадратического

отклонения шума, вычисляемая во всей области кадра. Результаты численного моделирования показывают, что практически оценивание σ^2 сводится к нахождению второго начального момента, так как математическое ожидание остаточного шума, как правило, близко к нулю. На изображениях, формируемых датчиками ИК диапазона, интересующие воздушные объекты зачастую имеют положительный тепловой контраст относительно фона, поэтому целесообразно подвергать пороговой обработке разность $d(i, j)$.

После получения $\hat{r}(i, j)$ бывает необходимо уменьшить степень фрагментации связанных областей бинарного изображения и подавить точечный шум, для чего может применяться морфологическая обработка бинарных изображений, заключающаяся в последовательном применении операций морфологического закрытия и открытия со структурирующими элементами заданных размеров. Измерение параметров объектов с использованием бинарных масок осуществляется с помощью известной процедуры разметки и параметризации описанной, например, в [1]. Для сопоставления списков объектов, найденных на текущем и предыдущем кадрах, и присвоения каждому объекту уникального идентификационного номера применяется алгоритм стробирования. Селекция объектов в стробе осуществляется по результатам анализа расстояния между центром объекта и центром строба. Размеры строба выбираются исходя из скорости движения и размеров объектов на изображении.

С целью повышения эффективности сопровождения объектов, размеры которых изменяются от кадра к кадру, предлагается ввести в структуру алгоритма дополнительный канал пространственной фильтрации и пороговой обработки. В обоих каналах осуществляется пространственная обработка по формулам (8)-(10), но значения q_2 , положение и размеры зоны обработки отличаются [10]. Каждый канал генерирует свое бинарное изображение, а для получения результата выделения объектов в виде единого бинарного изображения точки бинарного изображения 1-го канала, попадающие в область обработки второго канала, замещаются точками изображения с выхода канала 2. Первоначально работает только канал, предназначенный для выделения и обнаружения объектов, размеры которых лежат в широком диапазоне. Параметры канала фиксированы и не меняются с течением времени. При выделении интересующего объекта последний может быть взят на сопровождение. В этом случае включается второй канал, зона пространственной обработки которого соответствует стробу сопровождения интересующего объекта, что позволяет точнее оценить дисперсию шума на этапе получения бинарного изображения. Размер q_{2c} маски фильтра w_2 на следующем после включения дополнительного канала кадре выбирается исходя из выражения:

$$q_{2c}(n+1) = \begin{cases} q_{2c}(n)k_{cl} + \max(\tilde{V}_{об}, \tilde{W}_{об})(1-k_{cl}), \\ \text{при } \max(\tilde{V}_{об}, \tilde{W}_{об}) < q_{cp}, \\ q_{2c}(n)k_{cl} + 3\max(\tilde{V}_{об}, \tilde{W}_{об})(1-k_{cl})/2, \\ \text{иначе,} \end{cases}$$
(11)

где $(\tilde{V}_{об}, \tilde{W}_{об})$ – измеренные значения высоты и ширины объекта на изображении, k_{cl} – коэффициент сглаживания, q_{cp} – граничное значение q_{2c} .

Экспериментальные исследования

Экспериментальная проверка алгоритма проводилась на натуральных тестовых видеопоследовательностях ТВ и ИК диапазонов продолжительностью от 1500 до 2000 кадров. На рис. 3 приведены кадры из нескольких тестовых сюжетов. Фон на использованных сюжетах был как неподвижным, так и движущимся. Выбранные сюжеты характеризовались наличием как однородного, так и облачного фонов, размеры объектов лежали в диапазоне от 2x2 до 40x50 пикселей, отношение сигнал/шум достигало значения 2,5 и менее.



Рис. 3. Кадры из тестовых видеосюжетов (стрелками отмечено положение объекта)

Для обрабатываемой видеопоследовательности на каждом кадре находилось евклидово расстояние между эталонными координатами центра объекта $(\alpha_{об}, \beta_{об})$ и координатами центра m -го объекта $(\hat{\alpha}_{об}^{(m)}, \hat{\beta}_{об}^{(m)})$, найденного в обрабатываемой области кадра

$$d_{об}^{(m)} = \sqrt{(\hat{\alpha}_{об}^{(m)} - \alpha_{об})^2 + (\hat{\beta}_{об}^{(m)} - \beta_{об})^2}, \quad m = \overline{1, N_{об}}, \quad (12)$$

где $N_{об}$ – количество найденных объектов. После чего формировался бинарный параметр $N_{обн}^{(n)}$ из условия:

$$N_{обн}^{(n)} = \begin{cases} 1, & \text{если } \min_m d_{об}^{(m)} \leq 1 + \frac{1}{2} \sqrt{V_{об}^2 + W_{об}^2}, \\ 0, & \text{иначе.} \end{cases} \quad (13)$$

Объект считался обнаруженным в кадре n , если значение $N_{обн}^{(n)}$ было равно единице и потерянным в противном случае. Вычислялась частота правильного обнаружения объекта как отношение количества кадров, где объект был обнаружен, к длине видеопоследовательности. Ошибка определения координат центра объекта по

каждой из осей рассчитывалась как

$$(\Delta\alpha_{об}, \Delta\beta_{об}) = (\hat{\alpha}_{об}^{(s)} - \alpha_{об}, \hat{\beta}_{об}^{(s)} - \beta_{об}),$$

где $s = \arg \min_m d_{об}^{(m)}$.

Считая $\Delta\alpha_{об}$ и $\Delta\beta_{об}$ в случае правильного обнаружения независимыми случайными величинами, распределенными по гауссовскому закону с нулевым средним, можно оценить их СКО $\hat{\sigma}_{\Delta\alpha}, \hat{\sigma}_{\Delta\beta}$. При проведении экспериментальных исследований дополнительно находилась частота пропусков объектов. Под пропуском объекта понималась ситуация, когда объект не был обнаружен в течение заданного числа подряд идущих кадров (выбиралось равным двум). В таблице 1 приведены количественные характеристики эффективности работы рассматриваемого алгоритма (I), усредненные по нескольким тестовым сюжетам. В качестве примера в данной таблице приведены результаты работы байесовского сегментатора (II), описанного в [11].

Таблица 1.

Количественные характеристики эффективности работы алгоритмов, усредненные по нескольким тестовым сюжетам

Усредненная частота правильного обнаружения		Усредненная частота пропусков объекта		Усредненное СКО $(\hat{\sigma}_{\Delta\alpha}, \hat{\sigma}_{\Delta\beta})$, пикс.	
I	II	I	II	I	II
0,991	0,986	0,029	0,062	(0,67, 0,86)	(1,17, 1,88)

Анализ результатов показывает, что использование разработанного алгоритма выделения позволяет добиться уменьшения частоты пропусков объектов в среднем в 2 раза и СКО ошибки измерения координат центров объекта на 50-55% по сравнению с алгоритмом, основанным на байесовской сегментации. Предлагаемый подход также обладает невысокой вычислительной сложностью, позволяющей его реализовать на современной элементной базе. Так, количество операций сложения и умножения, затрачиваемых на обработку полного кадра при типовых параметрах одноканального алгоритма выделения объектов $q_1=3$, $q_2=15$, составляет $1,8 \times 10^7$.

С целью демонстрации работы алгоритма выделения воздушных объектов в реальных условиях наблюдения проводились его натурные испытания. Исследовалась эффективность сопровождения таких объектов как БПЛА, Ан-2, Миг-31, Ми-8, Ка-52 на дальностях до 30 км с помощью различных датчиков изображений. В целом, результаты проведенных испытаний свидетельствуют о возможности использования разработанного алгоритма в составе систем автоматического обнаружения и сопровождения объектов.

Заключение

Предложенный алгоритм способен выделять как малоразмерные, так и крупноразмерные объекты на облачных фонах при отношениях сигнал/шум до 2,5 в исходном изображении, обладает невысокой вычислительной сложностью, что позволяет обрабатывать полный кадр изображения в реальном масштабе времени. Разработанный подход может быть рекомендован для применения в составе систем автоматического сопровождения объектов. Исследования выполнены при поддержке гранта для ведущих научных школ НШ-1.2010.10.

Литература

1. Алпатов Б.А., Бабаян П.В., Балашов О.Е., Степашкин А.И. Методы автоматического обнаружения и сопровождения объектов. Обработка изображений и управление – М.: Радиотехника, 2008.–176 с.
2. Pham Q.H., Brosnan T.M., Smith M.J.T., Mersereau R.M. A morphological method of clutter suppression in ATR // Proc. of SPIE Vol. 3371, Automatic Target Recognition VIII. – 1998, pp. 367–374.

3. Алпатов Б.А., Бабаян П.В. Выделение движущихся объектов в условиях геометрических искажений изображений // Цифровая обработка сигналов.– 2004.– №4.– С. 9-14.
4. Liu Ya, Ai Haizhou, Xu Guangyou. Moving object detection and tracking based on background subtraction // Proc. of SPIE, Vol. 4554, Object Detection, Classification and Tracking Technologies, 2001, pp.62–66.
5. Amir H. Fazlollahi, Bahram Javidi Optimum receivers for pattern recognition problems with nonoverlapping target and background noise // Optical Engineering 36, 1997, pp.2633-2641.
6. Suyog D. Deshpande, M. H. Er, V. Rondab, Phillip Chan. Max-mean and max-median filters for detection of small-targets // Proc. of SPIE Vol. 3809, Signal and Data Processing of Small Targets, 1999, pp.74-83.
7. Алпатов Б.А. Оценивание параметров движущегося фрагмента двумерного изображения // Техника средств связи. Серия Техника телевидения.– 1991.– №2. С.77-81.
8. Стротов В.В. Оценивание параметров смещения изображения в задачах выделения движущихся объектов // Вестник РГРТУ.– Рязань, 2008.– №23.– С.30-37.
9. Муравьев В.С. Пространственный алгоритм обнаружения и измерения координат воздушных объектов на изображении // Вестник РГРТУ.– 2009.– №28.– С.17-20.
10. Муравьев В.С., Муравьев С.И. Алгоритм выделения и измерения координат объектов, наблюдаемых на облачных фонах // Вестник РГРТУ. – Рязань, 2007. – №21. – С. 20-24.
11. Бакут П.А., Лабунец В.Г. Телевизионная следящая система с байесовским дискриминатором цели // Зарубежная радиоэлектроника.– 1987.– №10.– С.81-93.

IMAGE PROCESSING ALGORITHM FOR AUTOMATIC AERIAL OBJECT TRACKING SYSTEMS

The extraction of aerial objects observed on television and infrared images and its role in the process of the automatic tracking is described. A typical structure of the real-time automatic tracking system is also presented. A model based object extraction algorithm suggested in the paper is served as a fundament for the modified approach. The results of an experimental research confirm high reliability characteristics of the proposed algorithm.

СЛЕЖЕНИЕ ЗА ОБЪЕКТОМ НА ОСНОВЕ АЛГОРИТМА СОПОСТАВЛЕНИЯ С ЭТАЛОНОМ ПРИ ОДНОВРЕМЕННОМ НАБЛЮДЕНИИ В ВИДИМОМ И ИНФРАКРАСНОМ ДИАПАЗОНАХ

Бабаян П.В., к.т.н., начальник лаборатории Рязанского государственного радиотехнического университета, e-mail: aitu@rsreu.ru

Смирнов С.А., аспирант Рязанского государственного радиотехнического университета, e-mail: aitu@rsreu.ru

Ключевые слова: слежение, оценка работоспособности, мультиспектральное наблюдение, комплексирование.

Введение

Технологии автоматического слежения за объектами в реальном времени широко используются для решения таких задач, как контроль транспортных потоков, поиск и спасение в чрезвычайных ситуациях, контроль общественного порядка, охрана важнейших объектов, навигация роботов. В настоящее время для автоматического слежения за объектами всё чаще используются мультиспектральные датчики, фиксирующие изображения наблюдаемой сцены в нескольких спектральных поддиапазонах инфракрасного, видимого или ультрафиолетового участков спектра.

По отношению к односпектральному наблюдению одновременное использование нескольких спектральных каналов, очевидно, позволяет повысить надежность слежения за объектами. Это объясняется, в частности, тем, что неблагоприятные условия наблюдения редко проявляются сразу во всех спектральных каналах. Например, недостаточный уровень естественного освещения сцены сказывается лишь на работе канала видимого диапазона, а слабый тепловой контраст объекта является проблемой лишь при наблюдении в инфракрасном диапазоне.

Очевидно, что разработка методов совместной обработки мультиспектральных изображений, позволяющих компенсировать недостатки одного спектрального канала за счет другого, открывает возможности для создания надёжных систем слежения, способных работать на больших дальностях, в любое время суток и в условиях слабой видимости. В настоящей статье предложен и исследован один из таких методов.

Постановка задачи

Пусть видеонаблюдение производится в двух спектральных диапазонах, и датчик формирует пары геометрически совмещенных изображений сцены, имеющих одинаковый размер. В поле зрения датчика изображений перемещается один объект. Примерами характерных объектов являются автомобили, самолеты, вертолеты, суда, люди. Фон, на котором наблюдается объект, может представлять собой облачное или ясное небо, пересеченную местность, городской ландшафт, водную поверхность. В процессе наблюдения могут происходить изменения условий наблюдения, связан-

Рассмотрена задача слежения за объектом в видеопоследовательности при одновременном наблюдении в видимом и инфракрасном диапазонах. Предложен алгоритм слежения, основанный на независимой обработке изображений в спектральных каналах. Выбор спектрального канала, обеспечивающего наилучшее качество слежения, производится на основе анализа признаков работоспособности. Представлены результаты экспериментальных исследований разработанного алгоритма.

ные с перемещением объекта на фон другого типа, изменениями освещенности, кратковременными частичными или полными заслонениями объекта фоном. В момент начала слежения поступает целеуказание, которое представляет собой параметры прямоугольника, описанного вокруг объекта. Целеуказание может быть сформировано оператором или получено техническими средствами.

Задача слежения состоит в том, чтобы на каждом кадре видеопоследовательности получить оценку текущих координат объекта.

В качестве базового подхода, на основе которого в настоящей работе строится алгоритм слежения, выбран подход на основе сопоставления с эталоном [1,2]. Известно, что при работе в одном спектральном диапазоне в описанных условиях наблюдения данный подход характеризуется высокой эффективностью и быстродействием. Таким образом, вопросы выбора базового подхода в дальнейшем не рассматриваются.

Опишем кратко основные особенности алгоритма сопоставления с эталоном, применяемого в настоящей работе. Алгоритм основан на поиске минимума разностной критериальной функции вида:

$$F(\alpha, \beta, n) = \sum_{(i, j) \in H} |l(i + \alpha, j + \beta, n) - h(i, j, n)|, \quad (1)$$

где n – номер кадра, $l(i, j, n)$ – наблюдаемое изображение, $h(i, j, n)$ – эталонное изображение объекта, H – множество точек изображения объекта. В результате минимизации (1) по (α, β) определяются измерения координат объекта $(\alpha^*(n), \beta^*(n))$, после чего эталонное изображение обновляется в соответствии с выражением:

$$h(i, j, n+1) = kh(i, j, n) + (1-k)l(i + \alpha^*(n), j + \beta^*(n), n), \quad (i, j) \in H, \quad (2)$$

где k – коэффициент ($0 < k < 1$).

Обновление эталонного изображения направлено на адаптацию к изменениям изображения объекта, происходящим в процессе наблюдения. Путем временной фильтрации измерений $(\alpha^*(n), \beta^*(n))$ могут быть получены оценки координат объекта в последовательности изображений.

Описанный алгоритм слежения ориентирован на работу в одном спектральном канале. В настоящей работе на его основе будет построен алгоритм, эффективный при использовании двух спектральных каналов наблюдения.

Общий подход к решению задачи

В литературе описаны несколько концепций совместного использования спектральных каналов наблюдения для слежения за объектами. Некоторые авторы предлагают решить задачу слежения отдельно в каждом канале, а затем объединить полученные результаты [3]. Вариацией этого подхода является объединение результатов слежения с использованием аппарата нечеткой логики [4]. В других публикациях описываются методы на основе отбора спектральных каналов, обеспечивающих наиболее успешное решение задачи слежения. В частности, в [5] предлагается использовать один канал в качестве основного источника информации для слежения за объектами, в то время как другие участвуют в процедуре адаптации параметров алгоритма. Наконец, в [6] представлен подход на основе совместной обработки изображений, соответствующих спектральным каналам. Отметим, что данные подходы относятся к категории методов комплексирования информации.

В данной работе выбрана концепция на основе раздельного слежения в каждом спектральном канале, при этом по окончании обработки очередной пары изображений выполняется оценка работоспособности алгоритма слежения для каждого канала. В качестве текущего измерения координат объекта берется измерение, полученное из канала, обладающего более высоким показателем работоспособности. Полученное измерение используется также для коррекции процесса слежения в канале, имеющем меньший показатель работоспособности.

Структура системы слежения, реализующей описанный подход, представлена на рисунке 1. Преимущество выбранной методики перед методиками, основанными на совместной обработке изображений, связано с относительной простотой построения мультиспектральной системы слежения на базе двух односпектральных. В то же время, для организации совместной обработки мультиспектральных изображений требуется достаточно много затрат в части разработки нового программного и аппаратного обеспечения. В отличие от известных подходов, в данной работе для оценки работоспособности используется признак, связанный с особенностями алгоритма слежения, а не с характеристиками изображения.

Оценивание ожидаемой работоспособности алгоритмов слежения за объектами.

Основной проблемой при реализации выбранного подхода является проблема оценки работоспособности алгоритма слежения, основанного на сопоставлении с эталоном. Для этого целесообразно использовать признак работоспособности (ПР), исследованный в работе [7] применительно к задаче выбора алгоритма измерения координат объекта. ПР вычисляется в соответствии с выражением

$$K(n) = \frac{\min_{(\alpha, \beta) \in \Gamma} F(\alpha, \beta, n-1)}{\min_{(\alpha, \beta)} F(\alpha, \beta, n)}, \quad (3)$$

где Γ – множество граничных точек зоны поиска.

Целесообразность применения данного признака вытекает из следующих рассуждений. Изображение объекта в последовательности кадров изменяется. Степень временной изменчивости объекта можно оценить минимальным значением критериальной функции. Если степень временной изменчивости объекта примерно соответствует степени отличия объекта и окружающего его фона, то измерение координат вряд ли возможно. В то же время, если объект не схож с фоном, и степень отличия объекта от фона будет больше, чем степень временной изменчивости объекта, то, скорее всего, измерение координат будет успешным.

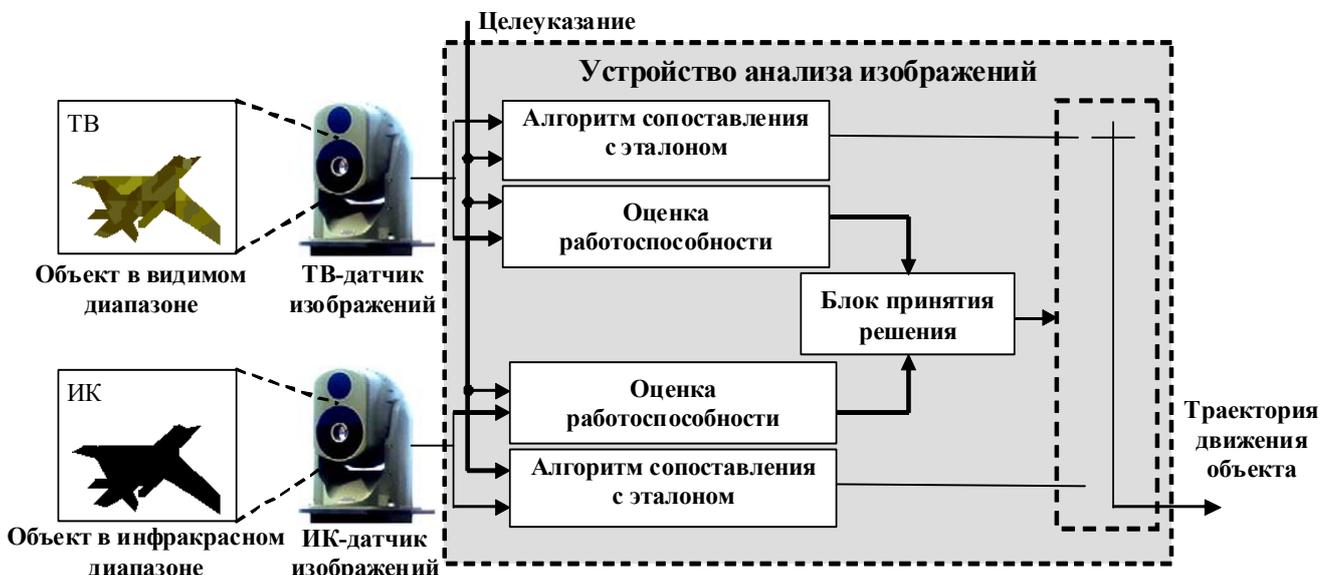
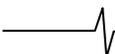


Рисунок 1 – Структурная схема системы слежения за объектами при наблюдении в ТВ и ИК диапазонах



Степень отличия изображения объекта от окружающего фона можно характеризовать величиной

$$\min_{(\alpha, \beta) \in \Gamma} F(\alpha, \beta, n-1).$$

Степень временной изменчивости изображения объекта описывается величиной

$$\min_{(\alpha, \beta)} F(\alpha, \beta, n).$$

Отношение этих величин можно использовать как признак, оценивающий работоспособность алгоритма слежения в n -м кадре.

Важным достоинством ПР вида (3) является возможность его использования для принятия решения о возможном срыве слежения. Для этого ПР сравнивается с порогом работоспособности T [7].

Для вычисления ПР не требуются эталонные данные о координатах объекта, однако ПР дает лишь оценочную, не всегда адекватную характеристику работоспособности. Несмотря на это, экспериментальные исследования, выполненные на большом количестве сюжетов, показали значительную корреляцию между данным ПР и качеством слежения за объектом.

Коррекция процесса слежения в канале, имеющем меньший показатель работоспособности

Представляется целесообразным рассмотреть два варианта такой коррекции:

- перезапуск слежения, при котором для целеуказания используются координаты, полученные в наиболее работоспособном канале;
- замена текущих координат объекта слежения на координаты, полученные в наиболее работоспособном канале.

Недостатком первого способа коррекции является потеря информации об изображении объекта при его кратковременном заслонении. Недостаток второго способа проявляется в условиях длительного пропадания

или заслонения объекта. В этом случае эталонное изображение объекта может быть искажено в результате сглаживания по формуле (2).

Экспериментальные исследования

Для проведения экспериментальных исследований использовалась тестовая база данных, включающая 32 пары видеосюжетов длиной от 90 до 120 кадров. Каждая пара состоит из сюжетов, снятых в ТВ и ИК диапазонах и сопровождается эталонными данными о траектории и линейных размерах объекта. Эталонные данные сформированы человеком-экспертом.

В роли критерия оценки эффективности того или иного алгоритма слежения использовалось количество сюжетов, на которых наблюдался срыв слежения, при этом срыв слежения фиксируется, если измеренные координаты объекта выходят за границы прямоугольной области, заданной эталонными данными о расположении и размерах объекта.

В таблице 1 приведены результаты экспериментальных исследований, полученные в односпектральном и двухспектральном вариантах слежения. Анализ результатов показывает, что предлагаемый подход позволяет улучшить качество слежения за счет использования двух спектральных каналов вместо одного, при этом наилучшим способом коррекции процесса слежения является замена текущих координат объекта слежения на координаты, полученные в наиболее работоспособном канале.

На рисунке 2 для одной из пар видеосюжетов представлены графики изменения ПР в каждом спектральном канале. До 40 кадра и начиная с 57 кадра слежение производится в ИК-канале. С 40 по 56 кадр происходят переключения между каналами. Начиная с 59 кадра ПР для ТВ канала опускается ниже порога работоспособности T .

Таблица 1

Результаты экспериментальных исследований

Алгоритм	Количество сюжетов, на которых происходит срыв слежения	
	ТВ-канал	ИК-канал
Один спектральный канал	8	3
Алгоритм с перезапуском слежения	2	
Алгоритм с заменой координат объекта	1	

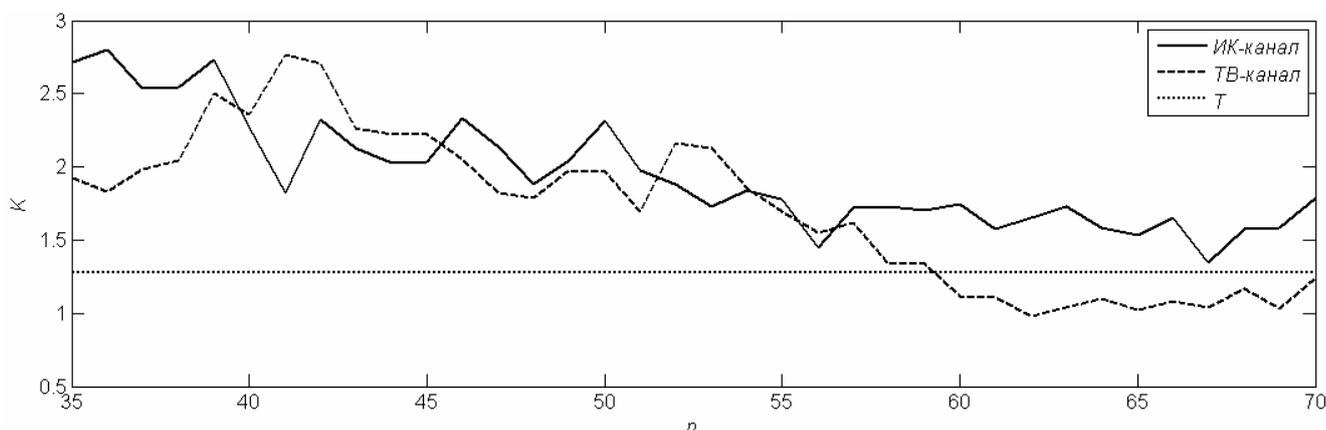


Рисунок 2. Признаки работоспособности для ИК-канала и ТВ-канала

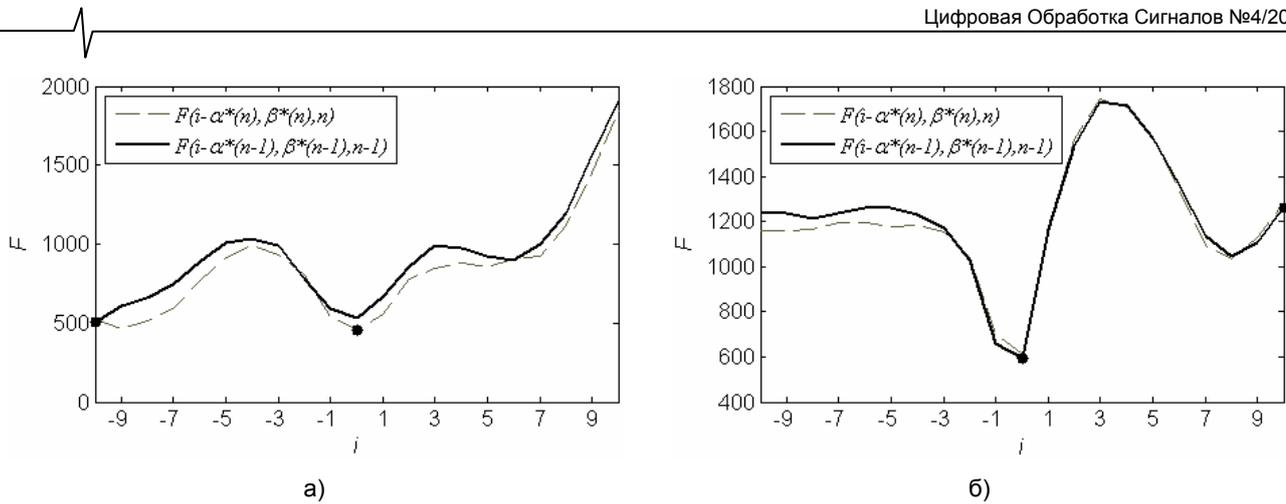


Рисунок 3. Поведение критериальных функций на кадрах N и $N-1$, соответствующих а) малому значению ПР ($K(n) \approx 1$); б) большому значению ПР ($K(n) \approx 2$)

На рисунке 3 показан характер изменения значений критериальных функций при различном значении ПР. Показаны соответствующие срезы значений критериальных функций. Рисунок 3,а соответствует ситуации, когда в зоне поиска присутствует фрагмент, похожий на объект, поэтому на кадре N минимальное значение критериальной функции приблизительно равно граничному минимуму критериальной функции из кадра $N-1$. Рисунок 3,б соответствует ситуации, когда похожие фрагменты отсутствуют, и как следствие, граничный минимум критериальной функции на кадре $N-1$ приблизительно в два раза больше минимума на кадре N . Точками отмечены значения критериальной функции, участвующие в вычислении ПР.

ЗАКЛЮЧЕНИЕ. Результаты экспериментальных исследований позволяют сделать выводы об эффективности предложенного подхода. Его существенным преимуществом также является возможность построения двухспектральной системы слежения без существенной переработки созданного ранее программного и аппаратного обеспечения. Отметим, что подход может быть распространен на случай большего числа спектральных диапазонов и большего числа базовых алгоритмов слежения.

Работа проводилась при финансовой поддержке Министерства образования и науки РФ (госконтракт № 14.740.11.0284).

Литература

1. Методы автоматического обнаружения и сопровождения объектов. Обработка изображений и управление / Алпатов Б.А., Бабаян 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. D. Borghys, P. Verlinde, C. Pemeel, and M. Achery, Long range

target detection in a cluttered environment using multi-sensor image sequences // Proceedings on Signal Processing, Sensor Fusion and Target Recognition IV - SPIE, 20-25 April 1997.

4. Jun Shen, Motion detection in color image sequence and shadow elimination // Proc. of SPIE. Vol. 5308, Visual Communications and Image Processing 2004, January 2004, pp. 731-740.
5. Vidya Manian, Luis O. Jimenez, Land cover and benthic habitat classification using texture features from hyperspectral and multispectral images // Journal of Electronic Imaging 16(2), 023011 (Apr–Jun 2007), pp. 1-12.
6. Алпатов Б.А., Бабаян П.В. Выделение движущихся объектов в последовательности мультиспектральных изображений при наличии геометрических искажений // Вестник Рязанского государственного радиотехнического университета. — Рязань, 2008. — №23. — С. 37-43.
7. Алпатов Б.А., Бабаян П.В., Смирнов С.А. Автоматическое слежение за объектами при отсутствии априорных сведений о фоновой обстановке // Цифровая обработка сигналов. – 2009. – №3. – С. 52-56.

OBJECT TRACING BY TEMPLATE MATCHING ALGORITHM AT SIMULTANEOUS OBSERVATION WITH VISIBLE AND INFRARED SENSORS

P.V. Babayan, S.A. Smirnov

In this paper the problem of object tracing in video sequence at simultaneous observation with visible and infrared sensors is described. The algorithm of tracing based on independent image processing in spectral channels is offered. The choice of spectral channel providing the best quality of tracking, is produced on the basis of the analysis of the performance indicator. The results of experimental research of the proposed algorithm are presented.

КОРРЕКЦИЯ АРТЕФАКТОВ ЗВОНА ПРИ JPEG-КОДИРОВАНИИ ИЗОБРАЖЕНИЙ

Бехтин Ю.С., д-р техн. наук, профессор кафедры автоматики и информационных технологий в управлении Рязанского государственного радиотехнического университета, e-mail: yuri.bekhtin@yandex.ru

Толстая Е.В., ст. инженер отдела трёхмерных приложений, Исследовательский центр Samsung Electronics, г. Москва; заочный аспирант кафедры автоматики и информационных технологий в управлении Рязанского государственного радиотехнического университета, e-mail: ktolstaya@yandex.ru

Ключевые слова: компрессия цифровых изображений, сжатие изображений с потерями, метод JPEG, артефакты звона, коррекция артефактов звона

Введение

Метод кодирования цифровых изображений JPEG широко применяется в цифровой фотографии и для хранения и передачи изображений в Интернет. При больших степенях сжатия метод JPEG вносит заметные артефакты в декодированное изображение. Принято различать на JPEG-изображениях блочные артефакты (blockiness) и артефакты звона (ringing), вызванные эффектом Гиббса [1].

Точная процедура обнаружения артефактов в каждом блоке изображения позволяет производить аккуратную модификацию изображения с подавлением артефактов и сохранением мелких деталей без внесения дополнительных погрешностей в изображение. Предлагаемая процедура обнаружения артефактов использует только коэффициенты дискретно-косинусного преобразования (ДКП). Устранение артефактов на изображении производится только для тех блоков изображения, где были выявлены артефакты. При этом результаты анализа коэффициентов ДКП сохраняются в памяти, а собственно процедура устранения артефактов звона на изображении производится после его декодирования.

К настоящему времени предложено множество способов обнаружения артефактов звона, которые основываются на нахождении контуров объектов в изображении [2-5, 10]. Однако нахождение контуров объектов – это достаточно сложная вычислительная процедура, имеющая ряд неизвестных параметров. В данной работе производится выявление артефактов звона без выполнения поиска контуров объектов.

Описывается экономный метод обнаружения и коррекции артефактов звона, появляющихся при блочном кодировании изображений с применением ДКП и квантования коэффициентов преобразования за счет избирательной (адаптивной) фильтрации участков, на которых предполагается появление звона.

Обнаружение артефактов звона на JPEG-изображениях

В отличие от блочных, артефакты звона не проявляются на границах блоков. Тем не менее, артефакты звона становятся заметными, когда блок, содержащий ненулевые коэффициенты ДКП, соответствующие высокочастотным гармоникам (особенно блок, содержащий значительный перепад яркости), соседствует с прилегающим гладким блоком. В этом случае возле края проявляются паразитные колебания яркости, которые принадлежат тому же блоку, что и край.

Закономерности проявления артефактов звона могут быть охарактеризованы следующими правилами:

- текущий блок изображения (содержащий артефакты) представлен достаточно большим количеством ненулевых коэффициентов ДКП;
- по крайней мере, один из соседних четырёх блоков является достаточно гладким блоком (представлен малым количеством ненулевых коэффициентов ДКП);
- количество нулевых коэффициентов ДКП, которыми представлен текущий блок изображения, также достаточно велико, так как артефакты, содержащиеся в блоке изображения, – признак того, что блок был подвержен сжатию.

На рис. 1 представлены фрагменты известного изображения «Маяк» и его ненулевые коэффициенты ДКП (чёрные клетки). Как видно из рис.1, блок, подверженный артефактам, имеет «гладкого» соседа.

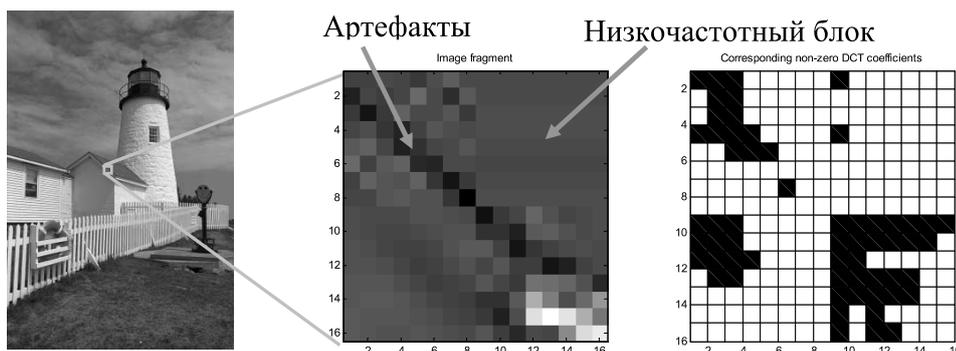


Рис.1. Фрагмент изображения «Маяк» (четыре соседних блока размером 8×8 пикселей), и соответствующие ненулевые коэффициенты ДКП (закрашены черным цветом)

На основе таких предположений проводится анализ коэффициентов ДКП для выявления «подозрительных» блоков. Методика расчета вероятности появления артефактов звона в блоке заключается в следующем.

Для каждого блока коэффициентов ДКП размером 8×8 рассчитываются следующие величины:

1. i_{\max} - максимальный индекс ненулевого ДКП коэффициента в упорядоченном векторе коэффициентов, полученном в результате зигзагообразного сканирования;
2. N_{DCT} - число ненулевых коэффициентов ДКП в блоке.

Относительно простая метрика для оценки вероятности появления артефактов RM рассчитывается по формуле:

$$RM = \frac{\sqrt{|G_1 G_2|} (64 - N_{DCT})}{64} \times \theta(N_{DCT} > 4) \cdot \theta(G_1 < 0) \cdot \theta(G_2 < 0) \cdot \theta(i_{\max} \geq 5), \quad (3)$$

где

$$G_1 = \min_{\substack{k=\text{"top"} \\ \text{"bottom"} \\ \text{"left"} \\ \text{"right"}}} \{N_{DCT}^k - N_{DCT}\}, \quad G_2 = \min_{\substack{k=\text{"top"} \\ \text{"bottom"} \\ \text{"left"} \\ \text{"right"}}} \{i_{\max}^k - i_{\max}\}, \quad (4)$$

$64 - N_{DCT}$ - число нулевых коэффициентов ДКП в текущем блоке. Чем больше это число, тем выше степень сжатия;

$\theta(G_1 < 0)$ – определяет ситуацию, когда рядом с текущим блоком находится, по крайней мере, один соседний блок с меньшим количеством ненулевых коэффициентов ДКП (т.е., по крайней мере, один из соседних четырех блоков является достаточно гладким блоком), при этом $\theta=1$, если $G_1 < 0$, и $\theta=0$ в противном случае;

$\theta(G_2 < 0)$ – определяет ситуацию, когда рядом с текущим блоком находится, по крайней мере, один соседний блок с меньшим индексом последнего ненулевого коэффициента ДКП (дополнительное условие наличия соседнего гладкого блока);

$\theta(i_{\max} \geq 5)$ – задает условие, что частота артефакта звона достаточно высока.

На рис.2 продемонстрировано два примера расчета величины RM. Те клетки, для которых вероятность RM

меньше 10 % ее максимального значения, были обнулены.

Представленная метрика позволяет спрогнозировать возможное местоположение артефактов (без использования ВЧ фильтра). Однако серьезным ее недостатком является большое количество ложных срабатываний (блок определен как имеющий артефакты, а на самом деле артефактов нет). Следующим шагом является определение силы проявляющихся артефактов.

Оценка яркости артефактов звона

К настоящему времени предложено множество метрик для оценки яркости артефактов звона, например, [2-4, 6, 7]. В большинстве своем расчет метрик производится путем определения областей перепада яркости и измерением осцилляций возле таких перепадов. К сожалению, подобные методы являются вычислительно сложными или требуют наличия эталонного (несжатого) изображения.

С помощью пороговой обработки метрики RM вида (3) удается, во-первых, снизить число потенциально «плохих» блоков и, во-вторых, оценить локальную дисперсию яркостей артефактов для последующей обработки изображения. Такая оценка позволяет использовать адаптивный фильтр для устранения выявленных артефактов.

Для оценки яркости артефактов звона недостаточно только вычислить дисперсию в блоке изображения, поскольку, скорее всего, блок будет содержать существенные перепады яркости из-за присутствия края объекта в блоке изображения. Для расчета дисперсии возле границы объектов необходимо сначала кластеризовать пиксели по яркости на несколько классов, затем для каждого класса рассчитать среднюю абсолютную разность между значениями пикселей и центром класса. На рис.3 показан декодированный блок яркостей пикселей с артефактами, содержащий край объекта. Пиксели кластеризуются по яркости на несколько классов [7], для каждого класса рассчитывается средняя абсолютная разность между значениями пикселей и центром класса, потом для оценки яркости артефакта выбирается меньшая величина из всех подсчитанных.

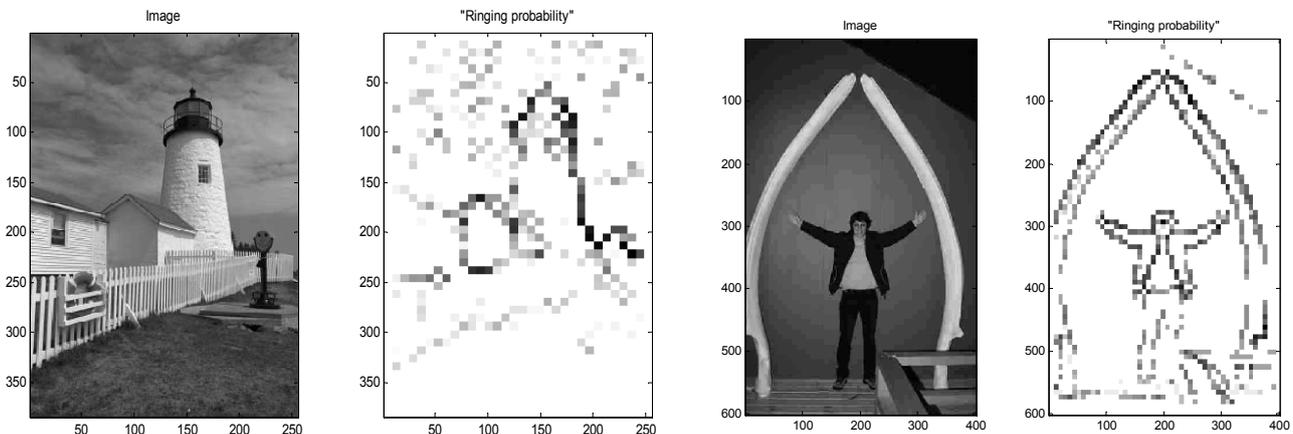


Рис.2. Пример двух изображений и их метрик RM вида (3)

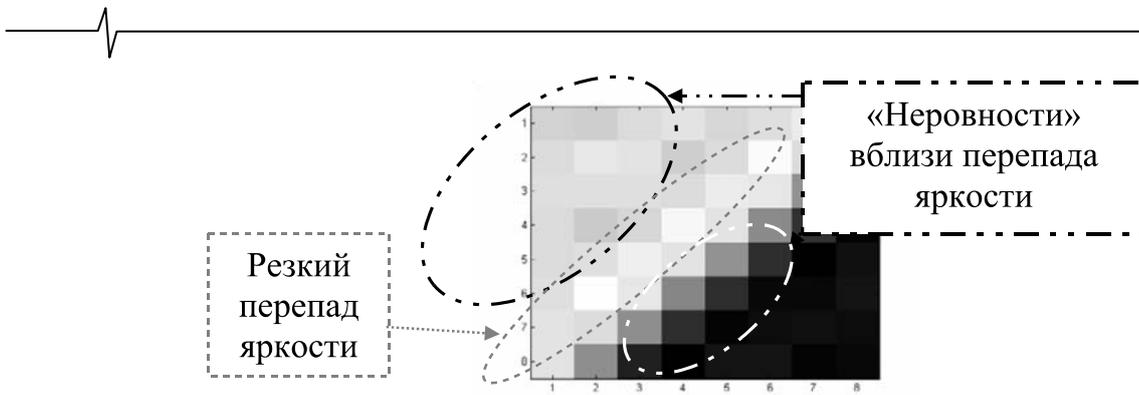


Рис.2. Декодированный блок, содержащий артефакты и резкий перепад яркости

В отличие от [7] предложено производить разделение на 2, 3 и 4 класса и выбирать наименьшую ошибку кластеризации. В результате с помощью вычисленной величины средней абсолютных разностей между значениями пикселей и центром класса можно оценить разброс значений пикселей внутри потенциально «испорченных» блоков изображения. Такая процедура относительно проста с точки зрения вычислительных и временных затрат. Вычисленная величина используется далее на этапе коррекции артефактов.

Коррекция артефактов звона

Артефакты звона эффективно корректируются с помощью билатерального фильтра [8]. При этом обрабатываются только те блоки, в которых оценка RM вида (3) больше нуля.

Для ускорения вычислений билатеральный фильтр применяется сначала для строк, затем для столбцов. Фильтрованное изображение I_f , таким образом, находится в два этапа:

$$I_f^{temp}(m, j) = \frac{\sum_{i=m-x}^{m+x} I(i, j) h(|I(i, j) - I(m, j)|) w(|i - m|)}{\sum_{i=m-x}^{m+x} h(|I(i, j) - I(m, j)|) w(|i - m|)}, \quad (5)$$

$$I_f(m, n) = \frac{\sum_{j=m-x}^{m+x} I_f^{temp}(i, j) h(|I_f^{temp}(m, j) - I_f^{temp}(m, n)|) w(|j - n|)}{\sum_{j=m-x}^{m+x} h(|I_f^{temp}(m, j) - I_f^{temp}(m, n)|) w(|j - n|)}, \quad (6)$$

где $h(y)$ - ядро фильтра по диапазону яркостей, $w(t)$ - ядро фильтра размером $2x+1$ в пространственной области:

$$h(y) = \exp\left(-\frac{y^2}{2\sigma_r}\right), \quad 0 < y < 255, \quad w(t) = \exp\left(-\frac{t^2}{2\sigma_d}\right), \quad -x < t < x, \quad (7)$$

здесь σ_r и σ_d - параметры ядра фильтра по диапазону яркостей и в пространственной области соответственно. Применение билатерального фильтра позволяет сгладить небольшие перепады яркости, возникающие вследствие артефактов, в то же время эта операция предохраняет значительные перепады яркости от размытия. Параметры σ_d и x могут быть зафиксированы (в экспериментах принималось $\sigma_d = 0,8$ и $x=7$), что существенно сокращает необходимые вычислительные ресурсы. Фильтрация по диапазону яркостей может быть адаптивной: параметр σ_r зависит от вычисленной на

предыдущем шаге величины средней абсолютной разности между значениями яркостей пикселей и центрами классов. Чем больше эта величина, тем больше значение параметра σ_r . Однако, как показали результаты статистического моделирования, подобная адаптивность незначительно улучшает качество декодированных изображений, поэтому для упрощения вычислений параметр σ_r является константой (в экспериментах принималось $\sigma_r = 0,1$). С другой стороны, чтобы избежать потери мелких деталей на декодированном изображении в случае, если артефакты звона не сильно заметны, предлагается зафиксировать σ_r и использовать величину средней абсолютных разностей между значениями яркостей пикселей и центрами классов при смешивании изображения и результата его фильтрации. Подобная операция замещения приводит к увеличению объема обработанного изображения в среднем на 3÷5 процентов.

Результаты статистического моделирования

Во многих работах для оценки качества алгоритма коррекции артефактов используются критерии минимума среднеквадратической ошибки (MSE) и максимума пикового отношения сигнал-шум (PSNR):

$$MSE = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N (I - I_{true})^2, \quad PSNR = 10 \log_{10} \left(\frac{255^2}{MSE} \right). \quad (8)$$

В статье приведены результаты сравнения по критерию максимума PSNR разработанного алгоритма с тремя известными приложениями для коррекции JPEG артефактов: shape-adaptive DCT (SA-DCT) [9], Topaz Labs DeJPEG (<http://www.topazlabs.com>), JPEG Enhancer (<http://www.vicman.net/jpegenhancer/>). В случае ручной корректировки использовались параметры по умолчанию. Результаты моделирования (обработки) для фрагмента изображения «Фотограф» показаны на рис.4-6.

Необходимо отметить, что, несмотря на объективность, критерий максимума PSNR является относительно грубым показателем качества, поскольку не всегда соответствует субъективному визуальному восприятию человека. Например, часто изображения с большими значениями PSNR могут казаться пользователю сильно размытыми.

Рис. 7. демонстрирует зависимость PSNR от степени сжатия, измеряемое через относительный показатель количества бит, приходящихся на один пиксель, - BPP (bit per pixel), для изображений «Фотограф» и «Лена». JPEG-изображения были получены в программной среде Photoshop CS2 с помощью функции «Save for Web».

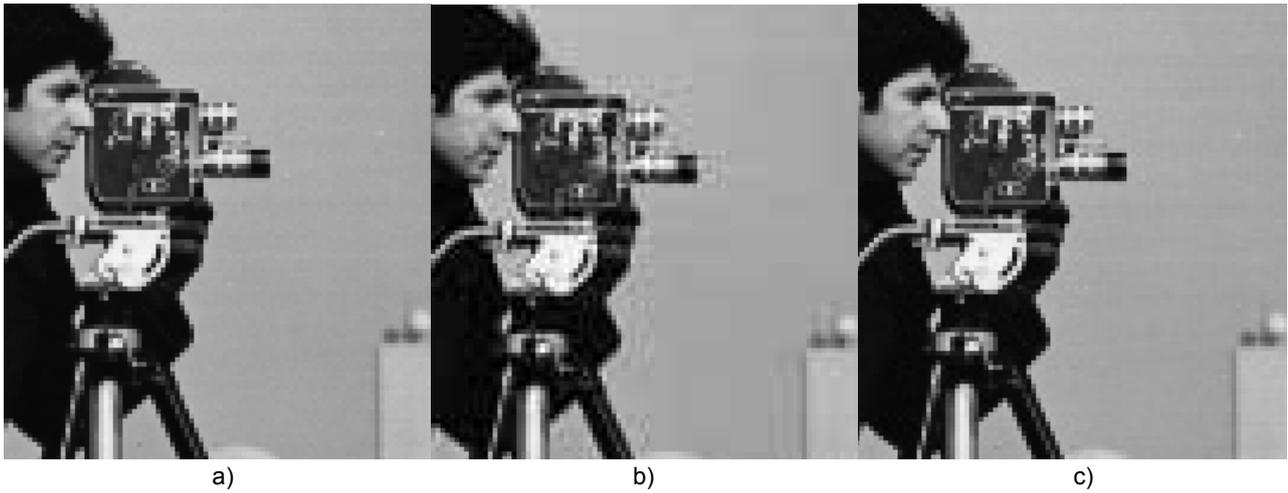


Рис.3. Фрагмент изображения «Фотограф»: а) исходное изображение (несжатое); б) 20%-е качество Photoshop (PS) “Save for Web”; в) 80%-е качество Photoshop “Save for Web”

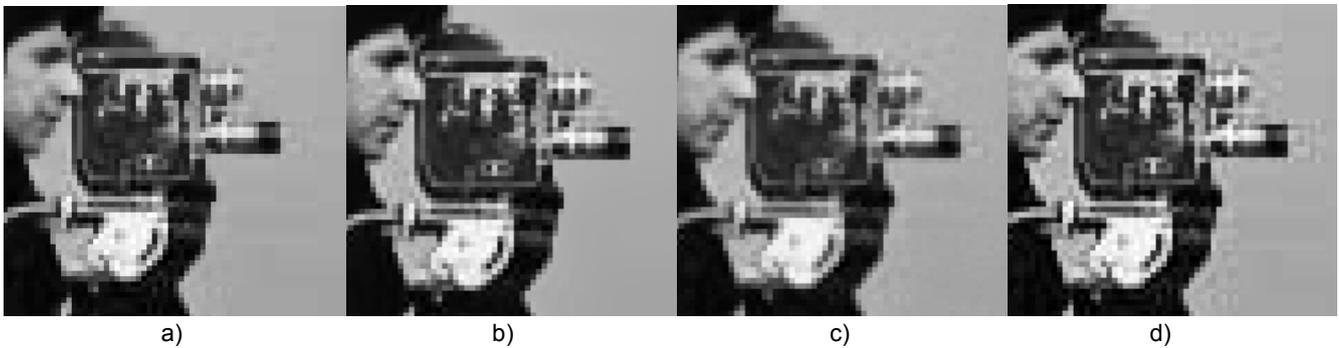


Рис.4. Результаты обработки изображения 20%-го качества Photoshop “Save for Web”: а) предложенный метод; б) Shape-adaptive DCT; в) JPEG Enhancer; г) DeJPG

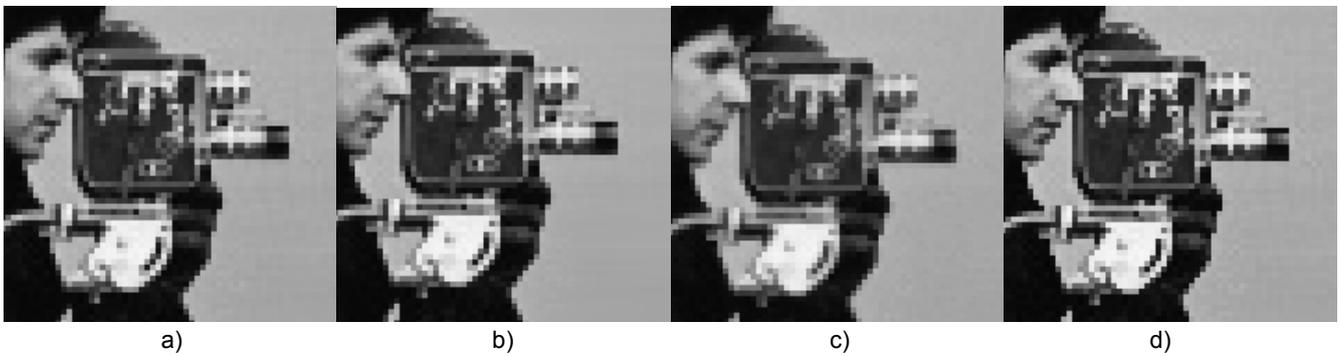
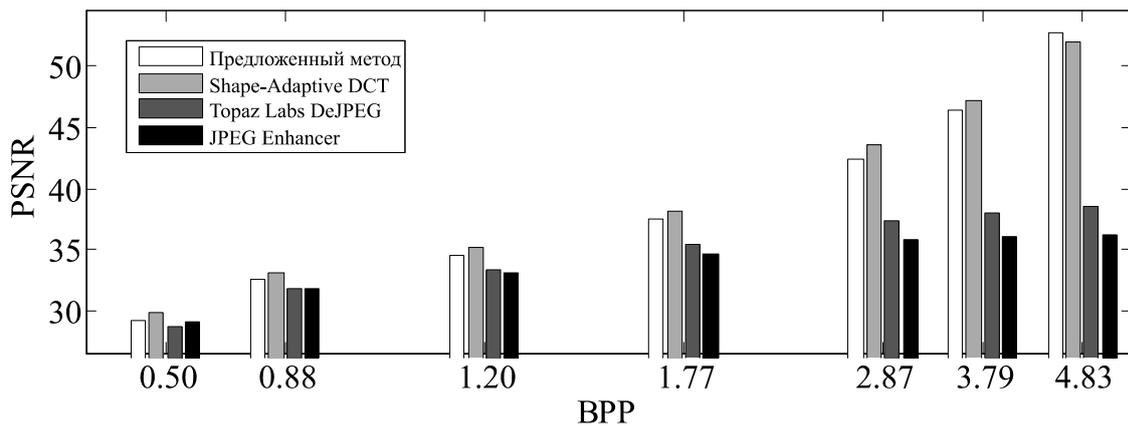


Рис.5. Результаты обработки изображения 80%-го качества Photoshop “Save for Web”: а) предложенный метод; б) Shape-adaptive DCT; в) JPEG Enhancer; г) DeJPG

PSNR для изображения "Фотограф"



PSNR для изображения "Лена"

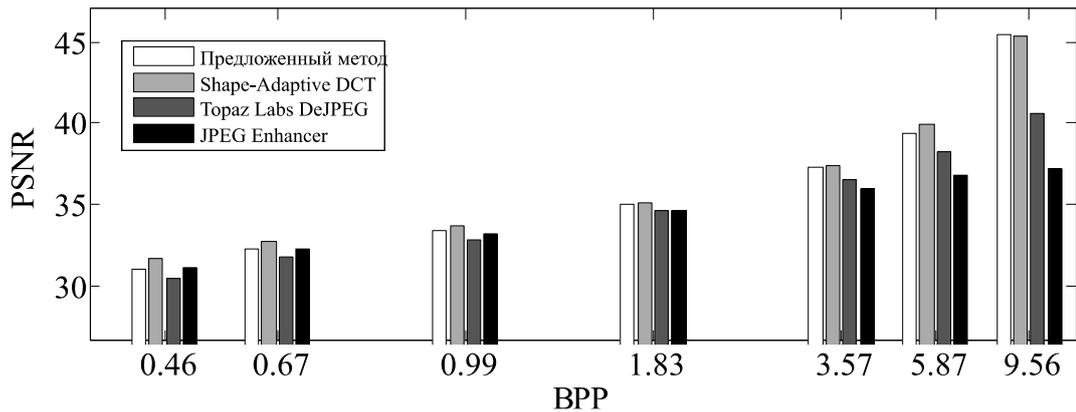
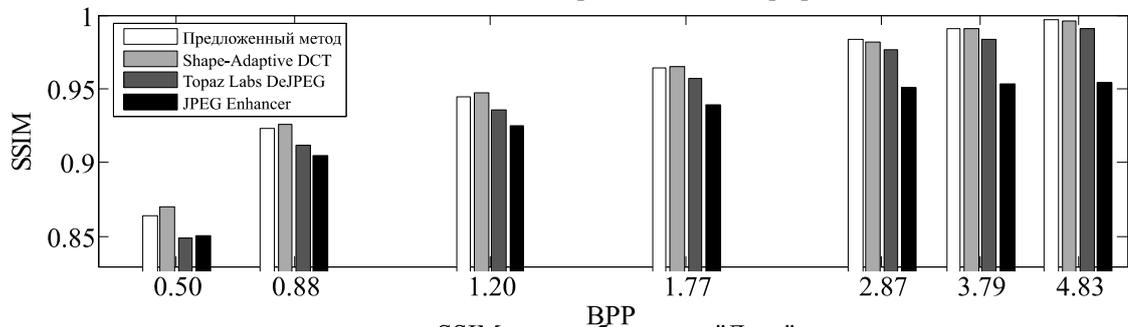


Рис.6. Результаты сравнения значений PSNR для изображений «Фотограф» (вверху) и «Лена» (внизу) для разных значений BPP

SSIM для изображения "Фотограф"



SSIM для изображения "Лена"

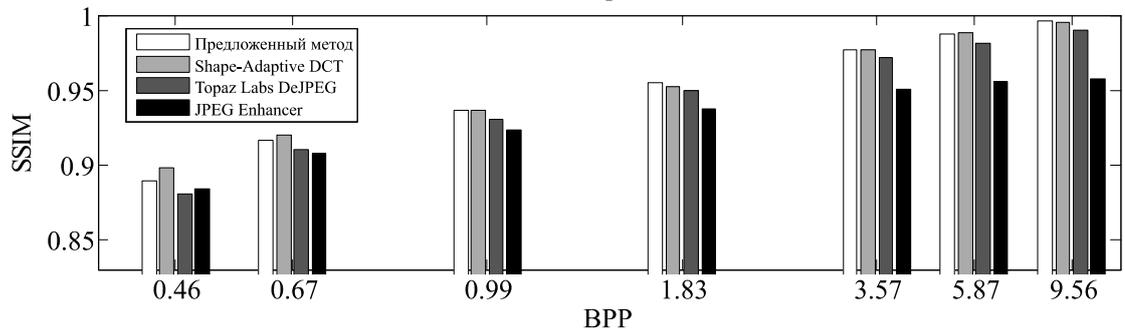


Рис. 7. Результаты сравнения значений SSIM для изображений «Фотограф» (вверху) и «Лена» (внизу) для разных значений BPP

При выявлении эффективности метода также использовалась комплексная оценка SSIM (Structural SIMilarity) [11]. Для экспериментов на ЭВМ была выбрана свободно распространяемая программа для среды Matlab [12], при этом применялись параметры «по умолчанию». Рис. 8 демонстрирует зависимость SSIM от степени сжатия BPP для тех же изображений.

Из рис. 7 видно, что предложенный метод немного уступает методу SA-DCT в смысле максимума PSNR при высоких (BPP<1) и средних степенях сжатия (1<BPP<3). Однако, по комплексной оценке SSIM, что иллюстрирует рис.8, предложенный метод не проигрывает методу SA-DCT при средних и высоких степенях сжатия. Вместе с

тем, при средних степенях сжатия изображения, обработанные методом SA-DCT, выглядят слишком размытыми, качество текстур снижено. Предложенный метод для изображений со средними и малыми (BPP>3) степенями сжатия (наиболее часто встречающийся случай) показывает относительно хорошие результаты: артефакты скорректированы, текстуры и края объектов не повреждены.

Кроме того, предложенный метод обеспечивает значительный выигрыш по быстродействию по сравнению с известными методами. В таблице приведено время обработки изображения размером 2480×3507 на ПК с процессором Pentium IV (3,2 ГГц, 1 Гб ОЗУ) различными методами.

Время обработки изображения размером 2480×3507 на ПК Pentium IV

Метод	Предложенный	UnJPEG	DeJPEG	JPEG Enhancer	SA-DCT (PS plug-in)
Время, с	9	26	37	25	811

Заключение

Разработан автоматический и вычислительно эффективный метод обнаружения и коррекции артефактов звона, появляющихся при JPEG-кодировании изображений. Метод основан на анализе коэффициентов ДКП и улучшает качество декодированных изображений при малых и средних степенях сжатия, по скорости обработки превосходя известные решения. В процессе обработки не используются сложные с вычислительной точки зрения операции, и не обязательно хранить все изображение целиком в памяти, что облегчает задачу интеграции соответствующих данному методу алгоритмов и программного обеспечения в различные устройства отображения, например, устройства фотопечати.

Литература

1. Адрес официального сайта для метода JPEG в Интернет: <http://www.jpeg.org>.
2. Feng, X., Allebach, J. P., "Measurement of ringing artifacts in JPEG images", Proc. SPIE 6076, pp. 74-83 (2006).
3. Oguz, S.H., Hu, Y.H., Nguyen, T.Q., "Image coding ringing artifact reduction using morphological post-filtering", IEEE 2nd Workshop on Mult.Sig.Proc., pp. 628-633 (1998).
4. Chang, Y.-W. and Chen, Y.-Y., "Alleviating-Ringing-Artifact Filter Using Voting Scheme", ICGST Int. Journal Graphics Vision and Image Proc.: Special Issue on Applicable Image Processing Techniques 6, pp. 9-14 (2006).
5. Marziliano, P., Dufaux, F., Winkler, S., Ebrahimi T. "Perceptual Blur and Ringing Metrics: Application to JPEG2000", Sig. Proc.: Image Comm. 19(2), pp. 163-172 (2004).
6. Chou, C., and Li, Y., "A perceptually tuned subband image coder based on the measure of just-noticeable distortion profile," IEEE Trans. Circuits Syst. Video Technol. 5(6), pp. 467-476 (1995).
7. Yang, S., Hu, Y., Nguyen, T. Q., and Tull, D. L., "Maximum-likelihood parameter estimation for image ringing-artifact removal," IEEE Trans. Circuits Syst. Video Technol. 11(8), pp. 963-973 (2001).
8. Tomasi, C., Manduchi, R., "Bilateral Filtering for Gray and Color Images," Proc. IEEE Conf. on Computer Vision, pp. 839-846 (1998).
9. Foi, A., Katkovnik, V., and Egiazarian, K., "Pointwise Shape-Adaptive DCT for High-Quality Denoising and Deblocking of Grayscale and Color Images," IEEE Trans. Image Proc. 16(5), pp. 1395-1411 (2007).
10. Vo, D. T., Nguyen, T. Q., Yea, S., Vetro, A., "Adaptive Fuzzy Filtering for Artifact Reduction in Compressed Images and Videos", IEEE Trans. Image Proc. 18(6), pp. 1166-1178 (2009).
11. Z. Wang, A. C. Bovik, H. R. Sheikh, Simoncelli, E. P., "Image quality assessment: From error visibility to structural similarity," IEEE Transactions on Image Processing, vol. 13, no. 4, pp. 600-612, Apr. 2004.
12. The SSIM Index for Image Quality Assessment, <http://www.ece.uwaterloo.ca/~z70wang/research/ssim/>
13. Толстая Е. В., Способ выявления искажений, вызванных эффектом Гиббса, при JPEG- кодировании, Патент РФ № 2365994 Опубликовано: 27.08.2009, Бюл. № 24.
14. Толстая Е. В. Способ удаления искажений на цифровых изображениях, Патент РФ № 2402070, Опубликовано: 20.10.2010 Бюл. № 29.

REMOVAL OF RINGING ARTIFACTS IN JPEG-CODED IMAGES

Yuri S. Bekhtin, Ekaterina V. Tolstaya

The paper is related to image quality enhancement; in particular, to the method for the effective reduction of ringing artifacts caused by lossy compression algorithms of JPEG coding using the block-based discrete cosine transform (DCT).

In order to reduce the effect of ringing artifacts due to significant information losses the variety of different algorithms and methods has been suggested in the last decades. However, the majority of solutions for ringing artifact detection usually relies on an edge-detection step that is a complicated and versatile procedure with unknown optimal parameters. In other cases, a reference (uncompressed) image is required to detect the ringing artifacts.

The presented method of ringing coding artifacts removal in images coded with block-based discrete cosine transform is characterized by next important properties: an edge-detection is not employed in proposed solution; the probability of the ringing artifact presence is computed in frequency domain using the simple and effective formula. The blocks having the high probability of ringing artifact presence are processed by the edge-preserving bilateral filter only; moreover, the introduced ringing artifacts severity metrics allow an end-user to vary the filter parameters.

The results of numerical modeling the algorithms of blocking and ringing artifact detection and removal are described and compared to some of the existing solutions in terms of PSNR and SSIM (Structural SIMilarity). The proposed method is intended to be implemented in software for digital cameras, slide scanners, camera-phones, printers, etc.

КОРРЕКЦИЯ СМАЗА ИЗОБРАЖЕНИЙ В СИСТЕМАХ КОСМИЧЕСКОГО НАБЛЮДЕНИЯ ЗЕМЛИ

Егошкин Н.А., к.т.н., с.н.с. НИИ «Фотон» РГРТУ, foton@rsreu.ru

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

Ключевые слова: фильтрация изображений, устранение смаза, приборы с зарядовой связью.

Постановка задачи

В современных космических системах высокодетального наблюдения Земли используются съемочные устройства на основе приборов с зарядовой связью (ПЗС), работающие в режиме временной задержки и накопления видеосигнала (ВЗН). В них развертка изображения в строчном направлении осуществляется за счет полета космического аппарата (КА), а в кадровом направлении – за счет опроса элементов ПЗС-матрицы, установленной перпендикулярно траектории полета (рис. 1). При съемке Земли в поле зрения строк с номерами $N, \dots, 2, 1$ последовательно попадают одни и те же объекты подстилающей поверхности. При каждом опросе ПЗС матрицы в течение времени τ осуществляется интегрирование лучистой энергии, после чего информация с $(i+1)$ -й строки переписывается в i -ю строку, а с первой ПЗС-строки снимается выходная информация.

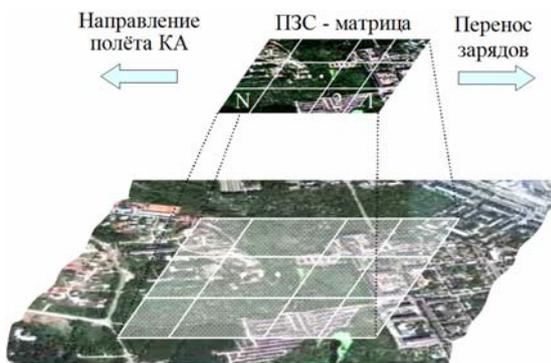


Рис. 1. Схема работы ПЗС-матрицы с ВЗН

Если за время периода T между опросами ПЗС-матрицы спутник смещается на расстояние Δ , и расстояние L между проекциями пикселей на Землю совпадает с этим смещением ($\Delta = L$), то каждая строка матрицы «видит» одни и те же объекты, и на выходе матрицы формируется сумма соответствующих сигналов. За счет этого достигается многократное увеличение отношения сигнал/шум [1].

Для правильной работы в режиме ВЗН необходимо, чтобы скорость движения КА была точно согласована с периодом опроса матрицы, т.е. $\Delta = L$. На практике равенство может нарушаться из-за ошибки вычисления

Рассматриваются модели смаза изображений при съемке поверхности Земли устройствами высокого разрешения, работающими на принципе временной задержки и накопления. Исследованы алгоритмы коррекции смаза, основанные на спектральном и алгебраическом подходах, выполнена их практическая апробация на снимках от космического аппарата «Ресурс-ДК».

скорости спутника и неточности задания частоты опроса ПЗС-матрицы. В результате возникают искажения типа «смаз» и, как следствие, стоит задача их коррекции. При этом необходимо выделить два вопроса. Во-первых, должны быть оценены с высокой точностью параметры смаза. Эта задача решается с использованием различных подходов, например [2], и в настоящей работе не рассматривается. Во-вторых, необходимо с использованием найденных параметров выполнить наиболее качественное устранение смаза, чему и посвящена настоящая статья.

Непрерывная модель смаза

Рассмотрим формирование сигнала одним произвольным столбцом ПЗС-матрицы. Пусть $s(x)$ – изображение подстилающей поверхности под этим столбцом, $\tau \ll T$, чувствительные элементы ПЗС-матрицы являются точками. Тогда на i -м периоде опроса в строке n накапливается заряд $q_n(i) = s(i\Delta + nL)$, $n = 1, N$, и на выходе рассматриваемого столбца матрицы формируется сигнал:

$$r(i) = \sum_{n=1}^N q_n(i-n) = \sum_{n=1}^N s((i-n)\Delta + nL) = \sum_{n=1}^N s(i\Delta + n(L-\Delta)). \quad (1)$$

Следовательно, процесс формирования сигнала может быть описан как свертка $s(x)$ с одномерной функцией рассеяния точки (ФРТ) вида

$$f(x) = \sum_{n=1}^N \delta(x - n(L-\Delta)), \quad (2)$$

с последующей дискретизацией непрерывного сигнала с шагом Δ , где $\delta(x)$ – дельта-функция. Выражение (1) можно представить в операторной форме как

$$r = D(f(s)) = Dfs, \quad (3)$$

где D – оператор дискретизации, f – оператор линейной системы (2).

В случае, когда время накопления τ и размеры чувствительного элемента конечны, (3) принимает вид

$r = D f g h s$, где h – оператор интегрирования сигнала по площади ПЗС-элемента, g – оператор смаза за время τ движущегося изображения. Оператор g действует как свертка с П-образной функцией рассеивания точки (ФРТ) шириной τ/T . Искажения, описываемые g и h , не связаны с процессом ВЗН, они возникают и при съемке Земли одной ПЗС-линейкой. В дальнейшем, не снижая общности, будем рассматривать модель искажений в виде (3), имея ввиду, что в (3) под f можно понимать композицию операторов f, g, h .

Дискретная модель смаза

На практике, как правило, работают с дискретными изображениями. Поэтому задачу устранения смаза можно сформулировать следующим образом: требуется перейти от дискретного изображения r к дискретному изображению r' , формируемому идеальной съемочной системой без смаза, для которой

$$r' = D s. \quad (4)$$

Изображение r' может содержать масштабные искажения с коэффициентом Δ/L , которые могут быть легко устранены при геометрической трансформации снимков.

Пусть изображение s таково, что может быть однозначно восстановлено по своим дискретным отчетам линейным оператором D^{-1} . Тогда $s = D^{-1} r'$, и после подстановки в (3), получим:

$$r = F r', \quad F = D f D^{-1}. \quad (5)$$

Оператор F связывает два дискретных изображения r и r' . Можно показать, что F линейен и инвариантен к целочисленным сдвигам входного дискретного сигнала r' . Следовательно, F действует как свертка с дискретной ФРТ, определяемой формулой (5). Далее без нарушения общности принимаем шаг дискретизации единичным.

На практике условия существования обратного оператора D^{-1} выполняются не точно. Поэтому в качестве D^{-1} следует использовать компактные интерполяционные операторы, например, линейные и кубические. Соответственно, возникает неопределенность в знаении F .

Определение N, Δ и L может быть выполнено различными методами, например, сопоставлением фактической частоты опроса ПЗС-матрицы со скоростью полета КА, или путем сопоставления изображений, синхронно формируемых в различных спектральных каналах [2]. В дальнейшем параметры N, Δ, L и ФРТ F будем считать известными.

Таким образом, задача устранения смаза сводится к решению уравнения (5) относительно r' , что может быть выполнено с использованием двух подходов: спектрального и алгебраического.

Спектральный подход по устранению смаза

В спектральной области связь искаженного и идеального изображений имеет вид $r = F r'$, где под r', r, F понимается преобразования Фурье соответствующих операторов из (5), а закон композиции есть поточечное умножение. Поэтому формально восстановление сигнала в спектральной области можно определить как $r' = r/F$.

Количественно смаз характеризует относительная ширина носителя функции (2), равная $\lambda = N |L - \Delta| / \Delta$. Приближенно под F можно понимать П-образную функцию с шириной λ , спектр которой $F_{\Pi}(\omega) = [\sin(0.5 \omega \lambda)] / (0.5 \omega \lambda)$. На рабочий участок спектра $\omega \in [0, \pi]$ приходится около 0.5λ нулей функции $F_{\Pi}(\omega)$.

Возможность восстановления изображения определяется видом функции $F_{\Pi}(\omega)$, в частности, числом нулей в рабочей полосе частот $\omega \in [0, \pi]$. При организации коррекции смаза следует различать три случая.

Незначительный смаз ($0 \leq \lambda < 2$). В данном случае нули $F(\omega)$ отсутствуют в рабочей полосе частот. Коррекция выполняется по формуле $r' = r/F$ и главная задача состоит в точном оценивании функций f и F , которые чаще всего не постоянны по полю снимка.

Средний смаз. Этому случаю соответствуют наличие несколько нулей $F(\omega)$. Здесь необходимо принимать меры по регуляризации функции $1/F$ вблизи нулей F .

Сильный смаз. Имеет место значительное число нулей $F(\omega)$ в рабочей полосе частот. В данном случае наблюдается сильное ослабление полезного сигнала, пропорциональное $1/(\omega \lambda)$, по мере роста частоты. Здесь необходимо управлять коэффициентом усиления $1/F$ на высоких частотах, иначе сигнал может быть полностью замаскирован шумом.

Оптимальное решение в смысле среднеквадратической близости исходного и восстановленного сигналов при некоррелированном шуме дает фильтр Винера [3]:

$$r'(\omega) = V(\omega) r(\omega), \quad (6)$$

$$V(\omega) = \frac{\tilde{F}(\omega)}{|F(\omega)|^2 + E(\omega)}, \quad (7)$$

где $E(\omega)$ – отношение спектральной плотности мощности шума и изображения. На практике точно оценить $E(\omega)$ трудно, поэтому в качестве $E(\omega)$ выбирается константа, которая подбирается оператором, исходя из наилучшего качества скорректированного изображения.

Недостатком фильтра Винера при коррекции смаза является искажение исходного изображения в виде теней от объектов на расстояниях, кратных величине смаза λ . Действительно, фильтрация по (6) и (7) равносильна прохождению эталонного сигнала $r(\omega)/F(\omega)$ через фильтр

$$K(\omega) = \frac{|F(\omega)|^2}{|F(\omega)|^2 + E(\omega)}. \quad (8)$$

Функция $K(\omega)$ действительная, а соответствующая ей импульсная характеристика $K(i)$ симметрична. На рис. 2 показан фрагмент $K(i)$ при $F(\omega) = F_{\Pi}(\omega)$, $\lambda = 100$ и $E(\omega) = 0.0001, 0.001, 0.01$. Как видим, амплитуда первой тени составляет около 1/3 от яркости объекта. Такие искажения хорошо заметны глазом.

Для уменьшения визуального восприятия искажений изображения тенями выполним модификацию оптимального фильтра (хотя при этом среднеквадратическая ошибка, конечно же, возрастет). Для этого составим целевую функцию

$$\varepsilon(K(\omega)) = \varepsilon_e(K(\omega)) + \sum_i \eta(i) K^2(i), \quad (9)$$

$$\varepsilon_s(K(\omega)) = (K(\omega) - 1)^2 + \frac{K(\omega)^2 E(\omega)}{|F(\omega)|^2}. \quad (10)$$

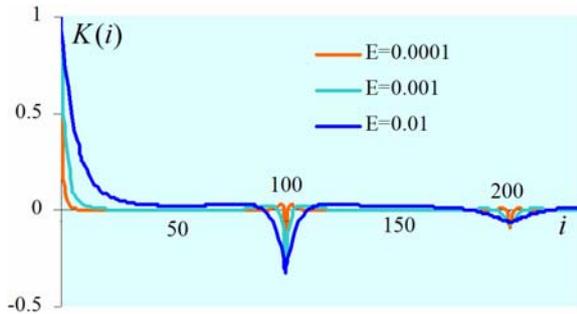


Рис. 2. Отклик фильтра $K(i)$ для П-образной ФРТ $F(\omega)$ при $\lambda = 100$.

Выражение (10) определяет среднеквадратическую ошибку, минимуму которой соответствует фильтр (7). В функции (9) второе слагаемое есть взвешенные с весами $\eta(i)$ отчеты импульсной характеристики. Задавая различные функции $\eta(i)$, например, отличные от нуля в районе выбросов, можно путем минимизации (9) синтезировать фильтры $K(\omega)$ с меньшей амплитудой выбросов.

На практике для минимизации (9) необходимо рассматривать конечные $K(i)$. При периодическом продолжении $K(i)$ спектр $K(\omega)$ становится дискретным и содержит конечное число отчетов. При этом, поскольку $K(i)$ линейно связана с отчетами $K(\omega)$, то минимум (9) легко находится по методу наименьших квадратов.

Минимизация (9) достаточно трудоемкий процесс. Более просто можно бороться с выбросами $K(i)$, регулируя поведение $V(\omega)$ вблизи нулей $F(\omega)$ при условии сохранения огибающей $V(\omega)$ такой же, как и у функции (7). Для этого представим восстанавливающий фильтр как

$$V(\omega) = \frac{\tilde{F}(\omega)}{|F(\omega)|^2 + \theta E(\omega)}, \quad (11)$$

где $O(\omega)$ – огибающая функции $F(\omega)$, θ – параметр. При $\theta = 1$ выражение (11) переходит в (7). При $\theta < 1$ добротность фильтра вблизи нулей $F(\omega)$ увеличивается и тени уменьшаются. При этом огибающая функции (11) сохраняется

такой же, как и у фильтра (7). Для функции $F(\omega) = [\sin(0.5\omega\lambda)] / (0.5\omega\lambda)$, можно принять $O(\omega) = 1 / (0.5\omega\lambda)$.

Следует отметить, что любые алгоритмы уменьшения теней приводят к увеличению ошибки (10), повышению шумов и увеличению длины ФРТ фильтра $F(\omega)$.

Практическую реализацию фильтрации в спектральной области удобно выполнять с использованием быстрого преобразования Фурье (БПФ). Для этого выбирается длина окна БПФ M из условия $M \gg \lambda$. По (7) или (11) рассчитывается дискретный спектр $V(\omega) = V(\omega)$

при $\omega = \frac{2\pi}{M} w$, $w = 0, M-1$. К спектру $V(\omega)$

применяется обратное БПФ и находится импульсная характеристика $V(i)$, $i = -M/2, M/2-1$. Далее анализируются «хвосты» $F(i)$ и, если они малы, то выполняется их усечение и соответствующее уменьшение размеров окна M . В противном случае M увеличивается для того, чтобы нивелировать эффекты циклической свертки, и процесс повторяется. В результате формируется длина окна M и характеристика восстанавливающего фильтра $V(i)$, $i = -M/2, M/2-1$. На рис. 3 показан фрагмент синтезированной $V(i)$ для П-образной F при $\lambda = 100$.

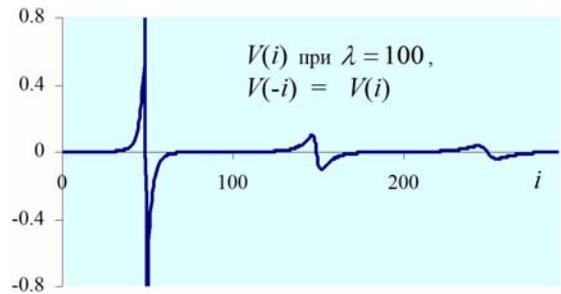


Рис. 3. Фрагмент ФРТ компенсирующего фильтра

Собственно восстановление изображения осуществляется его сверткой (не циклической) с фильтром $V(i)$, опять же с использованием БПФ. Для восстановления сигнала длиной в Z отчетов требуется $(Z + M)$ соседних отчетов входного сигнала.



Рис. 4. Коррекция смаза изображения на основе спектрального подхода

Рассмотренные выше модели, алгоритмы и технологии экспериментально исследовались с привлечением изображений от КА «Ресурс-ДК», в которых наблюдались сильные смазы из-за неправильного задания частоты опроса ПЗС-матриц. На рис. 4 слева представлены исходное смазанное изображение, справа – обработанное с использованием спектрального подхода. Оценки параметров смаза выполнялись в соответствии с технологиями, описанными в работе [2].

Алгебраический подход

Основан на составлении системы линейных алгебраических уравнений (СЛАУ), связывающих сигналы $r'(i)$ и $r(i)$, и её решении методами псевдообращения [3]. Методы эти трудоемки, т.е. практически применимы только к небольшим изображениям. Использование аппарата псевдообращения также не позволяет выявить физический смысл ограничений, за счет которых устраняется неоднозначность в решении СЛАУ. Для устранения этих недостатков предлагается следующее конструктивное решение.

Пренебрегая постоянным сдвигом по i , запишем связь идеального $r'(i)$ и искаженного $r(i)$ сигналов в виде свертки с причинным фильтром F :

$$r(i) = \sum_{j=0}^{\lambda} F(j) r'(i-j) \tag{12}$$

Тогда восстановление сигнала может быть выполнено рекурсивным фильтром:

$$r'(i) = \frac{r(i)}{F(0)} - \frac{1}{F(0)} \sum_{j=1}^{\lambda} F(j) r'(i-j) \tag{13}$$

Этот фильтр позволяет однозначно определить отчеты $r'(i)$, $i \geq 0$, по известным $r(i)$, $i \geq 0$, и начальным условиям $v(j) = r'(-j)$, $j = 1, \lambda$. Однако фильтр (13) не устойчив, поскольку влияние начальных условий не затухает со временем. Поэтому для применения (13) необходимо определять начальные условия, исходя, например, из минимума пульсаций выходного сигнала [4]:

$$\varepsilon(v) = \sum_{i=0}^M r'^2(i), \text{ где } M - \text{длина обрабатываемого фрагмента.}$$

Практическое решение уравнения $\varepsilon(v) = \min$ основано на линейности (13). Вначале находятся реакции $\zeta_j(i)$, $j = 1, \lambda$, фильтра (13) на λ линейно независимых начальных условий при входном сигнале $r(i) = 0$, а также реакция $\rho(i)$ фильтра (13) для $r(i) \neq 0$ и произвольных начальных условий. Далее составляется линейная комбинация

$$r'(i) = \rho(i) - \sum_{j=1}^{\lambda} a_j \zeta_j(i),$$

коэффициенты a_j которой находятся из условия

$$\sum_{i=0}^M r'^2(i) = \min$$

стандартным для евклидова пространства образом.

Особенно эффективно вычисление (13) осуществляется для П-образных ФРТ $F(j)$. В этом случае минимизация $\varepsilon(v)$ также упрощается, поскольку, как легко

показать, реакции $\zeta_j(i)$, $j = 1, \lambda$, есть всевозможные функции с периодом $\lambda + 1$.

На практике изображение обрабатывается пофрагментно, причем для каждого фрагмента начальные условия определяются отдельно. Важное значение имеет соотношение $k = M / \lambda$. Оно определяет остроту пиков фильтра, регуляризующего $1/F$ вблизи нулей F . Действительно, решение $\varepsilon(v) = \min$ равносильно прохождению точно восстановленного сигнала через линейный фильтр с ФРТ $K(i)$, срезающий некоторые сигналы. Форма $K(i)$ показывает, как искажается полезный сигнал.

На рис. 5 представлен график $K(i)$ для П-образной $F(j)$ при $\lambda = 99$, $k = 5$.

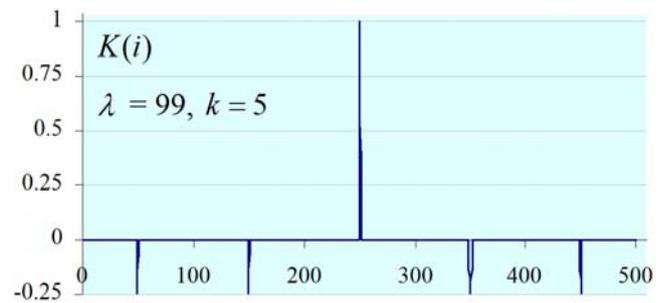


Рис. 5. Функция $K(i)$ при алгебраической фильтрации

Таким образом, и при алгебраической фильтрации возникают тени от объектов на расстояниях, кратных λ . Амплитуда теней составляет примерно $1/(k-1)$ от яркости объектов, тени имеют одинаковую амплитуду, число теней равно $(k-1)$. Увеличивая M и k , можно снижать уровень теней (но за счет роста их числа).

Рассмотренный подход позволяет устранить неустойчивость фильтра (13) и регулировать добротность множителя $K(i)$ путем изменения величины M . Однако для фильтрации сильного смаза этого не достаточно, поскольку сигнал маскируется шумом из-за подъема высоких частот фильтром (13). Для решения этой проблемы выходной сигнал (13) следует пропустить через фильтр низких частот (ФНЧ), например $K(\omega) = |O(\omega)|^2 / (|O(\omega)|^2 + E(\omega))$, уже встречавшийся в выражении (11). Точная реализация такого ФНЧ на практике едва ли оправдана, достаточно реализовать какую-либо аппроксимацию этого фильтра. Хорошие результаты дает, например, простейшая аппроксимация вида

$$K'(\omega) = \frac{1+2a}{1+ae^{i\omega} + ae^{-i\omega}}, \tag{14}$$

где $a = 0.5(K(\pi) - 1) / (K(\pi) + 1)$. Выражению (14) соответствует СЛАУ следующего вида:

$$r''(i) = (1+2a)r'(i) - ar''(i-1) - ar''(i+1), \tag{15}$$

где $r''(i)$ - искомый сигнал. СЛАУ (15) является трехдиагональной и легко решается с помощью алгоритма прогонки.

На рис. 6 представлены смазанное изображение КА «Ресурс-ДК» и результаты его обработки с использованием алгебраического подхода.

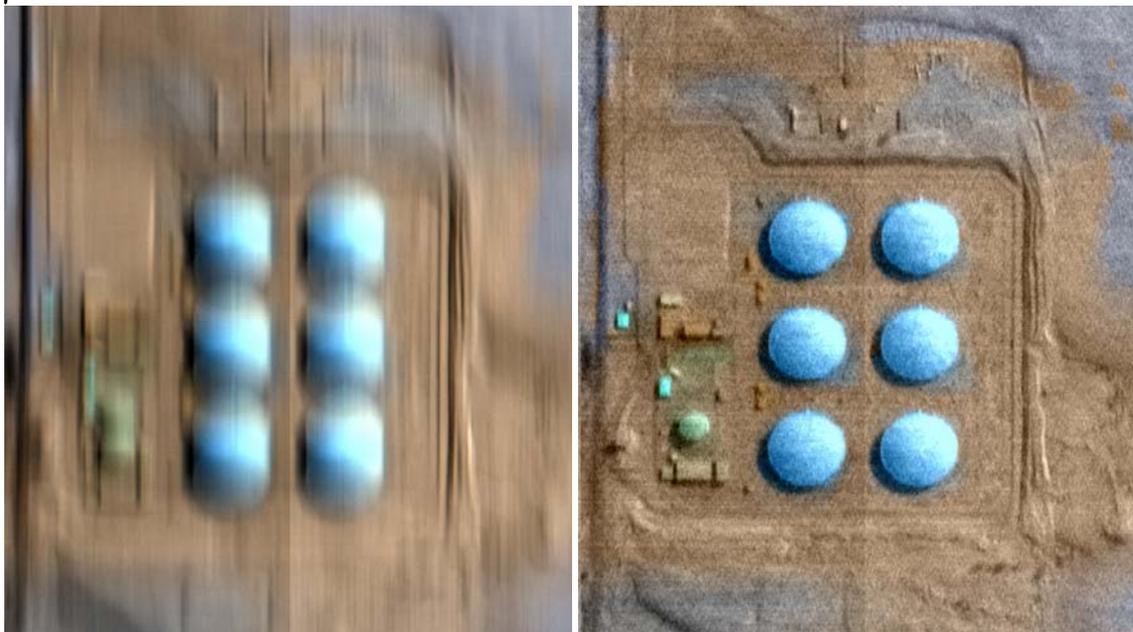


Рис. 6. Коррекция смаза изображения на основе алгебраического подхода

Заключение

По результатам выполненных исследований можно сделать следующие выводы.

1. Исследован спектральный подход к коррекции смаза. Его использование в виде классического фильтра Винера позволило достаточно качественно выполнить коррекцию смаза. Однако при таком решении рассматриваемой задачи появляются визуально заметные тени от объектов.

2. Выполнена модификация классического фильтра, основанная на регулировании его добротности. Это позволило ценой некоторого увеличения среднеквадратической ошибки восстановления добиться заметного уменьшения наблюдаемости теней.

3. Исследован алгебраический подход к коррекции смаза. На реальных космических изображениях установлено, что этот подход дает несколько худшие результаты при коррекции смаза. В тоже время предложены вычислительные процедуры, которые позволяют многократно повысить скорость решения рассматриваемой задачи.

Работа проводилась в рамках реализации ФЦП «Научные и научно-педагогические кадры инновационной России» на 2009 – 2013 годы (госконтракт № П2422).

Литература

1. Пресс Ф. П. Фоточувствительные приборы с зарядовой связью. М.: Радио и связь, 1991. 264 с.
2. Кузнецов П.К., Семавин В.И., Солодуха А.А. Алгоритм компенсации скорости смаза изображения подстилающей поверхности, получаемого при наблюдении Земли из космоса // Вестник Сам. ГТУ, 2005. №37. С. 150-157.
3. Грузман И.С., Киричук В.С. Цифровая обработка изображений в информационных системах. Новосибирск: Изд-во НГТУ, 2002. 352 с.
4. Еремеев В.В., Злобин В.К., Обработка аэрокосмических изображений. М.: Физматлит, 2006. 406 с.

CORRECTION SMEAR IMAGES IN THE SPACE SURVEILLANCE SYSTEM EARTH

Egoshkin N.A., Ereemeev V.V.

The models of image blurring when photographing the Earth's surface field device, high-resolution operating on the principle of time-delay and accumulation leniya. Studied algorithms for correcting blurring, based on spectral and algebraic-ray approaches, performed their practical testing on images from space vehicles is the "Resurs-DK".

УДК 004.415.2

ХЭШ-ТЕЙЛОВАЯ ПИРАМИДАЛЬНАЯ РЕОРГАНИЗАЦИЯ ИЗОБРАЖЕНИЙ В СИСТЕМАХ ОБРАБОТКИ ДАННЫХ ДИСТАНЦИОННОГО ЗОНДИРОВАНИЯ ЗЕМЛИ

Побаруев В.И., к.т.н., с.н.с., НИИ «Фотон» Рязанского государственного радиотехнического университета, тел. (4912) 46-03-72, e-mail: foton@rsreu.ru

Светелкин П.Н., к.т.н., с.н.с., НИИ «Фотон» Рязанского государственного радиотехнического университета, тел. (4912) 46-03-72, e-mail: foton@rsreu.ru

Ключевая слова: дистанционное зондирование земли, обработка данных, матрица изображений, пирамидальное перемешивание строк, организация памяти.

Рассматривается метод реорганизации матриц изображений в виртуальной памяти программных систем обработки данных дистанционного зондирования Земли, обеспечивающий высокую скорость при интерактивной обработке видеоданных сверхбольшого объема за счет пирамидального перемешивания данных.

Введение

Развитие средств космической съемки идет по пути повышения разрешения снимков и увеличения полосы захвата, что приводит к росту объемов изображений. Например, на космическом аппарате (КА) «Ресурс-ДК», установлено электронное съемочное устройство сканерного типа, выполняющее съемку земной поверхности в трех спектральных диапазонах. В результате съемки формируются спектрозональные изображения размером до 36000×300000 пикселей, соответственно, такой многозональный снимок может занимать более 60 Гбайт дискового пространства. Интерактивная обработка изображений таких объемов на современных персональных компьютерах требует разработки специальных методов видеоинформационного обмена.

Современные программные средства обработки изображений больших объемов решают проблему организации эффективного видеоинформационного обмена по-разному. Наиболее эффективным в настоящее время подходом является создание подсистемы виртуальной памяти. Собственная подсистема виртуальной памяти необходима для преодоления ограничений адресуемой памяти, например, в 32-разрядных операционных системах семейства Windows, а также для обеспечения более эффективного кэширования данных на основе априорных знаний о процессе обработки.

Наиболее простой является матричная организация виртуальной памяти, когда образ изображения в неизменном виде располагается в файле и отображается на оперативную память (либо загружается построчно). Однако матричная организация имеет главный недостаток: неоптимальная работа с памятью, диском и файлом подкачки при геометрической обработке, а также недостаточная скорость отображения изображений при уменьшении масштаба отображения.

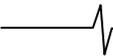
Для ускорения геометрической обработки наиболее эффективно использование тейловой организации видеоданных [1]. Подобная организация данных используется, например, в таких широко распространенных программных продуктах, как ERDAS, PHOTOMOD. Однако при тейловой организации доступ к сверхбольшим

изображениям при уменьшении масштаба просмотра также выполняется довольно медленно из-за необходимости чтения всего тейла данных в операции отображения, хотя реально требуются лишь отдельные байты тейла. Для ускорения операции отображения изображений на мониторе в уменьшенном масштабе, например, в программном комплексе ERDAS IMAGINE, используется пирамида изображений [2]. Но, как показали практические исследования, использование пирамидального представления изображений в памяти тоже имеет ряд недостатков. Во-первых, необходимо дополнительно хранить в памяти и на диске несколько пирамидальных уровней изображения, на создание которых затрачивается некоторое время. Во-вторых, при интерактивной обработке изображений необходима синхронизация между обработанным изображением и его пирамидальными представлениями, что требует дополнительных вычислительных затрат.

Описание метода

Для устранения указанных выше недостатков предлагается хэш-тейловый пирамидальный метод организации видеоданных, основанный на реорганизации матрицы изображения с целью обеспечения быстрого доступа к данным при различных масштабах отображения. Данный метод использует хэширование (перемешивание) строк матрицы с целью формирования пирамидальных образов без увеличения объема самой матрицы. При этом возможно перемешивание как по строкам, так и по столбцам одновременно. В настоящей работе для упрощения понимания метода рассматривается хэш-тейловая организация памяти с хэшированием только по строкам.

Для пояснения идеи предлагаемого метода рассмотрим вначале пример тейловой организации, который показан на рис. 1. Экспериментально размеры тейла установлены следующие: количество строк $Q_{cmp} = 16$, длина подстроки $L_{cmp} = 4096$ байт. Таким образом, первая строка изображения длиной 36000 пикселей (стандартная ширина кадра от КА «Ресурс-ДК») и радиометрическим разрешением 16 бит/пиксель располагается в первой строке тейлов 0-17. При этом количест-



во элементов в строке тейла зависит от радиометрического разрешения снимка и числа спектральных каналов. Обмен с оперативной и дисковой памятью производится всегда целыми тейлами.

Рассмотрим тейловую организацию памяти в сравнении с хэш-тейловой организацией. Применительно к рис. 1 расположение строк в тейлах при тейловой и хэш-тейловой организациях показано на рис. 2.

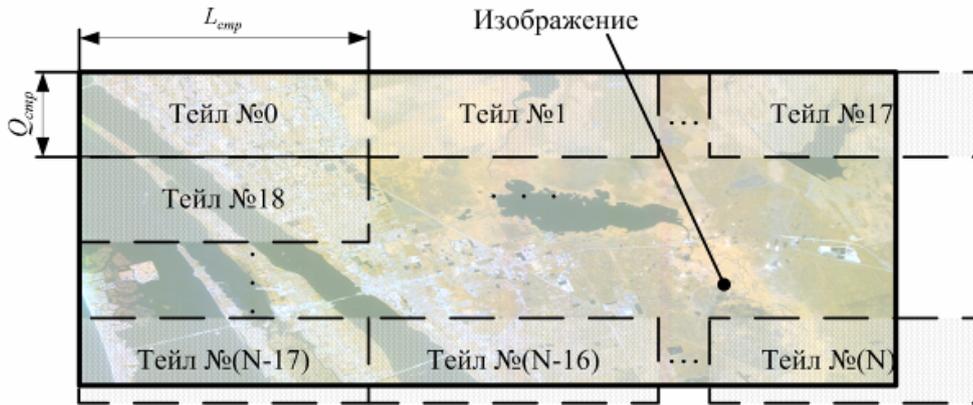


Рис. 1. Расположение изображения в памяти при тейловой организации

Номер строки изображения

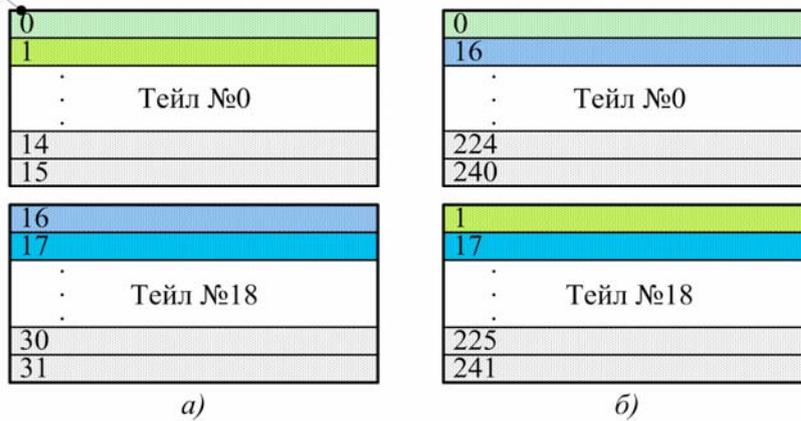


Рис. 2. Расположение строк изображения в тейлах: а – тейловая организация, б – хэш-тейловая организация.

Из рис. 2 видно, что при хэш-тейловой организации в тейлы помещаются строки с межстрочным интервалом $S = 16$, который подобран эмпирически. Такой подход обеспечивает последовательное расположение данных на диске организованных в виде прореженных образов изображений.

При считывании части строки изображения с номером n , начиная с номера пикселя m необходимо определить номер тейла, в котором находится строка и смещение начала части строки относительно начала тейла.

При тейловой организации номер тейловой строки в матрице тейлов определяется по формуле

$$NMT_{стр} = n / Q_{стр}, \quad (1)$$

где $NMT_{стр}$ – номер тейловой строки в матрице тейлов. Затем определяется номер тейла NT в матрице тейлов,

$$NT = NMT_{стр} \cdot QMT_{стб} + N_{noz} / Q_{стр}, \quad (2)$$

$$N_{noz} = m \cdot Q_{кан} \cdot Q_{бп},$$

где $QMT_{стб}$ – количество столбцов в матрице тейлов, N_{noz} – позиция пикселя в байтах, $Q_{кан}$ – количество каналов изображения, $Q_{бп}$ – разрядность изображения (байт/пиксель). Смещение $NT_{см}$ первого пикселя запрашиваемой строки n относительно начала тейла оп-

ределяется по формуле

$$NT_{см} = (n \bmod Q_{стр}) \cdot L_{стр} + (N_{noz} \bmod L_{стр}), \quad (3)$$

где \bmod – операция вычисления остатка от деления.

При хэш-тейловой организации номер тейловой строки в матрице тейлов вычисляется по формуле

$$NMT_{стр} = \frac{n}{Q_{стр} \cdot S} \cdot S + (n \bmod Q_{стр}). \quad (4)$$

Заметим, что в формуле (4) сокращать на S нельзя, т.к. все операции являются целочисленными. Номер тейла определяется по формуле (2). Смещение первого пикселя относительно начала тейла:

$$NT_{см} = \frac{n \bmod (Q_{стр} \cdot S)}{S} \cdot L_{стр} + (N_{noz} \bmod L_{стр}), \quad (5)$$

На основе приведенных формул были реализованы механизмы тейловой и хэш-тейловой организации, а также получены оценки временных затрат. Замеры времени проводились на персональном компьютере с процессором Pentium 4 3 ГГц, 2 Гбайт оперативной памяти (ОП). Для проведения первого теста (табл. 1) был установлен размер кэша приложения 1,3 Гбайт, количество строк изображения составляло 15000. Это необходимо для того, чтобы фрагмент считываемого изображения полностью помещался в кэш приложения.

Таблица 1

Временные затраты на считывание 15000 строк изображения

	Строчная организация		Тейловая организация		Хэш-тейловая организация	
	а)	б)	а)	б)	а)	б)
Последовательное считывание строк	18,72	0,05	18,56	1,19	18,42	1,17
Считывание каждой 16 строки	6,86	0,03	18,95	0,13	18,80	0,08

а – время первого считывания, с; б – время повторного считывания, с.

Таблица 2

Временные затраты на считывание 105000 строк изображения

	Строчная организация		Тейловая организация		Хэш-тейловая организация	
	а)	б)	а)	б)	а)	б)
Последовательное считывание строк	313,0	443,22	231,14	146,36	296,89	166,59
Считывание каждой 16 строки	119,80	51,72	238,50	146,67	161,67	12,45
Считывание каждой 64 строки	24,44	6,67	110,83	36,88	5,02	0,03

а – время первого считывания, с; б – время повторного считывания, с.

Из табл. 1 видно, что если изображение полностью помещается в кэш приложения, то время последовательного считывания с диска при строчной, тейловой и хэш-тейловой организациях практически одинаково. При считывании каждой 16 строки строчная организация выигрывает в 2,7 раза при первом считывании, это связано с предвыборкой данных с диска при считывании в уменьшенном масштабе. Видно, что при небольших размерах изображения строчная организация выигрывает у тейловой и хэш-тейловой организаций за счет более быстрой и простой программной реализации.

Учитывая, что целью настоящей работы является ускорение работы со сверхбольшими изображениями, необходимо получить оценки для изображения большего объема, например, с количеством строк 105000 (для тестов использовалось реальное изображение). Для достоверности этого теста необходимо, чтобы размер ОП, выделенной под кэш приложения, был более чем в 2 раза меньше, чем размер занимаемого изображением дискового пространства, деленный на $Q_{cmp} = 16$. То есть необходимо проанализировать скорость подгрузки хэш-тейловых блоков с диска в процессе отображения изображения с масштабом не менее 1:16. Так как тест выполнялся применительно к изображению размером 36000×105000 пикселей, то занимаемое дисковое пространство составляет примерно 7210 Мбайт. Отсюда, размер кэша приложения в ОП должен составлять менее чем $7210 / (16 \cdot 2) \approx 225$ Мбайт. Для выполнения теста (табл. 2) был установлен размер кэша приложения равный 219 МБ.

Как видно из табл. 2, при хэш-тейловой организации наблюдается ускорение при работе с файлом подкачки приложения. При последовательном считывании изображения, не помещающегося в кэш, хэш-тейловая организация проигрывает по времени тей-

ловой в среднем в 1,2 раза и незначительно выигрывает у строчной. Это связано с разницей в предвыборках данных с диска. При считывании каждой 16-й строки получаем выигрыш хэш-тейловой организации в 1,5 раза над тейловой организацией и проигрыш в 1,3 раза строчной организации при первом считывании. При повторном считывании каждой 16-й строки получаем выигрыш в 11,8 раза в сравнении с тейловой и в 4,2 раза в сравнении со строчной организациями. При дальнейшем уменьшении масштаба преимущество хэш-тейловой организации в скорости считывания данных увеличивается.

Вторым средством ускорения отрисовки изображения при смене масштаба является подмена строк. Без подмены строк при отрисовке изображения, например, в масштабе просмотра 1:16 считывается каждая 16-я строка, а при переходе к масштабу 1:19 – каждая 19-я строка, т.е. при изменении масштаба приходилось каждый раз считывать большие объемы данных. При использовании подмены строк в масштабе 1:19 отрисовывается не каждая 19-я строка, а берется ближайшая строка из масштаба с ближайшим меньшим номером, равным степени двойки k . Например, без подмены строк отрисовываются строки с номерами 0, 19, 38, 57, 76, 95 и т.д., при использовании подмены, исходя из того, что ближайший меньший к 1:19 масштаб 1:16 ($2^k = 2^4 = 16$), отрисовываются строки 0, 16, 32, 64, 80, 96 и т.д. Таким образом, новая информация при переходе от масштаба 1:16 к масштабу 1:19 практически не считывается, за счет этого и достигается высокая скорость отрисовки. В табл. 3 приведено сравнение объемов считываемой с диска информации при визуализации изображения на экране с разрешением 1280×1024 пикселей при смене масштаба 1:16 на 1:19.

Сравнение объемов считываемых с диска данных

	Объем считываемых данных	
	без подмены строк	с подменой строк
Масштаб 1:16	87,9 Мбайт (1280 строк по 70,3 Кбайт)	87,9 Мбайт (1280 строк по 70,3 Кбайт)
Переход к масштабу 1:19	87,9 Мбайт (1280 строк по 70,3 Кбайт)	16,5 Мбайт (дополнительно считываем 240 строк по 70,3 Кбайт)



Рис. 3. Фрагмент изображения в масштабе 1:16: а – без интерполяции, б – с интерполяцией

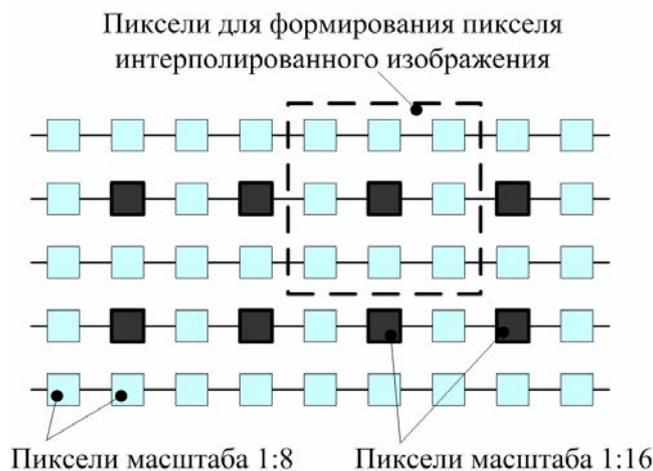


Рис. 4. Пример формирования интерполированного изображения при отрисовке в уменьшенном масштабе

Из табл. 3 видно, что за счет применения подмены строк, получаем более чем 5-кратное уменьшение объема считываемых данных при переходе от масштаба 1:16 к 1:19. При последовательной смене масштаба, т.е. при смене масштаба 1:16 на 1:17, получаем более чем 16-кратное уменьшение объема считываемых с диска данных, благодаря подмене строк по сравнению с обычным подходом.

Заметим, что при уменьшении масштаба просмотра изображения появляется «зернистость» (рис. 3, а), для ее устранения реализована интерполяция изображения (рис. 3, б). Для интерполяции берется масштаб уменьшения $1:2^{k-1}$ и отрисовывается средняя яркость по девяти ближайшим пикселям, как показано на рис. 4, т.е. при отрисовке в масштабе 1:16 ($1:2^k$) информация для яркостной интерполяции берется из масштаба 1:8 ($1:2^{k-1}$).

Заключение

Результаты практических испытаний показали, что предложенный метод хэш-тейловой пирамидальной организации видеоданных превосходит по быстродействию используемый в ERDAS IMAGINE. При этом не требуется дополнительных затрат дискового пространства и временных затрат на создание пирамиды.

Рассмотренный механизм хэш-тейлового пирамидального представления реализован в программных комплексах GeoScan, OrthoNormScan и других, разработанных в НИИ «Фотон» при РГПТУ и применяющихся для обработки изображений от КА «Ресурс-ДК». Практическое использование этого программного обеспечения получило высокие оценки со стороны потребителей.

Литература

1. JieBing Yu, David J. DeWitt. Processing Satellite Images on Tertiary Storage: A Study of the Impact of Tile Size on Performance. Proceedings of the 1996 NASA Conference on Mass Storage Systems, College Park, Maryland, September 1996.
2. ERDAS Field Guide. Fifth Edition, Revised and Expanded. 1999.

PYRAMIDAL HASH-ENCODING IMAGE MATRIX REORGANIZATION FOR**REMOTE SENSING DATA PROCESSING SYSTEMS***Pobaruev V.I., Svetelkin P.N.*

Method of image matrix reorganization of PC virtual memory for remote sensing data processing systems is consider. The goal of method is a high-water mark interactive performance of system by the use of pyramidal hash-encoding mass data images.



13-я Международная Конференция
**ЦИФРОВАЯ ОБРАБОТКА СИГНАЛОВ
 И ЕЕ ПРИМЕНЕНИЕ**

30 марта – 01 апреля 2011 г., Москва, Россия
The 13th International Conference **DIGITAL SIGNAL PROCESSING
 AND ITS APPLICATIONS**

ИНФОРМАЦИОННОЕ ПИСЬМО*Уважаемый коллега!*

*Приглашаем Вас принять участие в работе 13-й Международной конференции
 «Цифровая обработка сигналов и ее применение»,
 которая состоится в Москве 30 марта – 01 апреля 2011 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 (ЗАО МНИТИ)
- Компания AUTEX Ltd. (ЗАО «АВТЭКС»)

ПРИ УЧАСТИИ:

- Федеральное агентство по промышленности РФ
- Министерство образования и науки РФ
- ФГУП «ГРЧЦ»
- ГСКБ «АЛМАЗ-АНТЕЙ»
- ЗАО «Инструментальные системы»
- НТЦ «МОДУЛЬ»
- ЗАО «СКАН Инжиниринг Телеком»
- ГУП НПЦ «Элвис»
- Владимирский государственный университет
- Московский авиационный институт
- Московский государственный технический университет им. Н.Э. Баумана
- Московский институт радиотехники, электроники и автоматики
- Московский технический университет связи и информатики
- Московский энергетический институт
- Рязанский государственный радиотехнический университет
- Санкт-Петербургский государственный университет телекоммуникаций
- Санкт-Петербургский государственный электротехнический университет - ЛЭТИ
- Ульяновский государственный технический университет
- Ярославский Государственный Университет
- Московский Физико-технический институт (университет)

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
- Цифровое радиовещание
- ЦОС в системах защиты информации
- Проектирование и техническая реализация систем ЦОС
- ЦОС в открытых системах
- Проблемы подготовки специалистов в области ЦОС

По опыту прошедших «DSPA», оказалось очень эффективным местом общения для всех участников мероприятия. - представителей науки и производства, заказчиков и разработчиков, потребителей и поставщиков, преподавателей и студентов, работодателей и соискателей, научно-технических издательств и читателей.

РАБОТА КОНФЕРЕНЦИИ будет организована в форме пленарных и проблемно-тематических заседаний, стендовых докладов и заседаний «круглого стола». **В качестве пленарных** будут заслушаны заказные доклады по основным направлениям работы Конференции.

Рабочие языки конференции - **русский и английский.**

ОСОБЕННОСТИ АНАЛИЗА И ОБРАБОТКИ ИНФОРМАЦИИ ОТ СИСТЕМ ГИПЕРСПЕКТРАЛЬНОЙ СЪЕМКИ ЗЕМНОЙ ПОВЕРХНОСТИ

Антонушкина С.В., с.н.с. НИИ «Фотон» Рязанского государственного радиотехнического университета (РГРТУ), к.т.н., foton@rsreu.ru

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

Макаренков А.А., аспирант

Москвитин А.Э., с.н.с. НИИ «Фотон» РГРТУ, к.т.н.

Ключевые слова: гиперспектральные изображения, гиперкуб, анализ гиперспектральной информации, комплексирование, гистограмма.

Введение

В практику дистанционных исследований Земли в последние годы активно внедряются системы гиперспектральной съемки. Гиперспектральная съемка земной поверхности осуществляется в процессе движения спутника или самолета путем приема оптической системой излучения от Земли, его расщепления по длине волны и регистрации с помощью приборов с зарядовой связью (ПЗС-линеек), которые устанавливаются друг за другом в фокальной плоскости датчика, перпендикулярно полету носителя. В результате движения носителя и периодического опроса ПЗС-линеек формируется ГСИ, так называемый гиперкуб, в виде набора снимков, каждый из которых характеризует энергию излучения объектов наблюдаемой сцены в очень узком диапазоне спектра. По такому принципу построена, например, гиперспектральная аппаратура, установленная на космических аппаратах EarthObserver-1, Terra, Aqua и др. В России гиперспектральную аппаратуру впервые планируется установить на космический аппарат «Ресурс-П».

В результате расщепления входного потока излучения при формировании гиперспектральных снимков многократно повышается уровень шума по отношению к полезному сигналу, это создает основную трудность при анализе и обработке ГСИ. В тоже время ГСИ, в отличие от панхроматических и спектросональных снимков, предоставляют уникальную возможность более качественно решить задачу автоматической классификации на изображениях заданных объектов на основе анализа их спектральных характеристик.

В настоящей работе представлены первые результаты, полученные при решении задач сегментации и комплексирования ГСИ. Экспериментальная проверка предложенных решений выполнена с привлечением реальной информации, полученной в результате гиперспектральной аэросъемки Земли, которая предоставлена авторам ОАО «Красногорский завод им. С.А. Зверева» (рис. 1).

Меры сходства элементов ГСИ

На тех или иных мерах сходства элементов изображения основаны многие алгоритмы кластеризации и комплексирования. В отличие от традиционных мер, основанных на

Рассматриваются вопросы анализа гиперспектральных изображений (ГСИ) земной поверхности: меры сходства элементов ГСИ, комплексирование ГСИ со снимками от других съемочных систем, кластеризация объектов изображения. Представлены результаты исследований алгоритмов анализа гиперспектральных изображений, полученных в результате аэросъемки земной поверхности.

сопоставлении пространственно-энергетических характеристик, ГСИ дополнительно предоставляет для каждого элемента изображения детальную информацию о распределении энергии излучения по частоте (частотную характеристику). Это дает возможность вводить новые более качественные меры сходства элементов изображения.

Введем в рассмотрение гиперспектральное изображение $B \in \{B_{kmn}, k = \overline{1, K}, m = \overline{1, M}, n = \overline{1, N}\}$, где k – номер спектрального диапазона, (m, n) – координаты элемента изображения (его номер по строке и по столбцу), B – энергетическая яркость. Фиксированному k соответствует изображение, зарегистрированное в k -м спектральном диапазоне в виде матрицы значений энергетических яркостей: $B_{kmn}, m = \overline{1, M}, n = \overline{1, N}, k = const$. Точке изображения с фиксированными координатами (m, n) соответствует частотная характеристика в виде вектора из K значений энергетических яркостей для каждого частотного диапазона: $B_{kmn}, k = \overline{1, K}, m = const, n = const$.

Рассмотрим три частотно-пространственные меры сходства двух соседних по строке точек ГСИ – среднеквадратическую, корреляционную и спектрально-угловую [2], которые соответственно описываются следующими выражениями:

$$\Delta_{mn} = \sqrt{(B_{mn} - B_{m+1,n})^2}, \quad (1)$$

$$\rho_{mn} = \frac{B_{mn} \cdot B_{m+1,n} - \overline{B_{mn}} \cdot \overline{B_{m+1,n}}}{\sigma_{mn} \cdot \sigma_{m+1,n}}, \quad (2)$$

$$\alpha_{mn} = \arccos \left(\frac{B_{mn} \cdot B_{m+1,n}}{\sqrt{B_{mn}^2} \cdot \sqrt{B_{m+1,n}^2}} \right), \quad (3)$$

где σ_{mn} и $\sigma_{m+1,n}$ – среднеквадратические отклонения отсчетов спектральных характеристик для точек (m, n) и $(m+1, n)$, а черта над переменными и выражениями означает операцию усреднения по K спектральным диапазонам,

например: $\overline{B_{mn}} = \frac{1}{K} \sum_{k=1}^K B_{mnk}$.

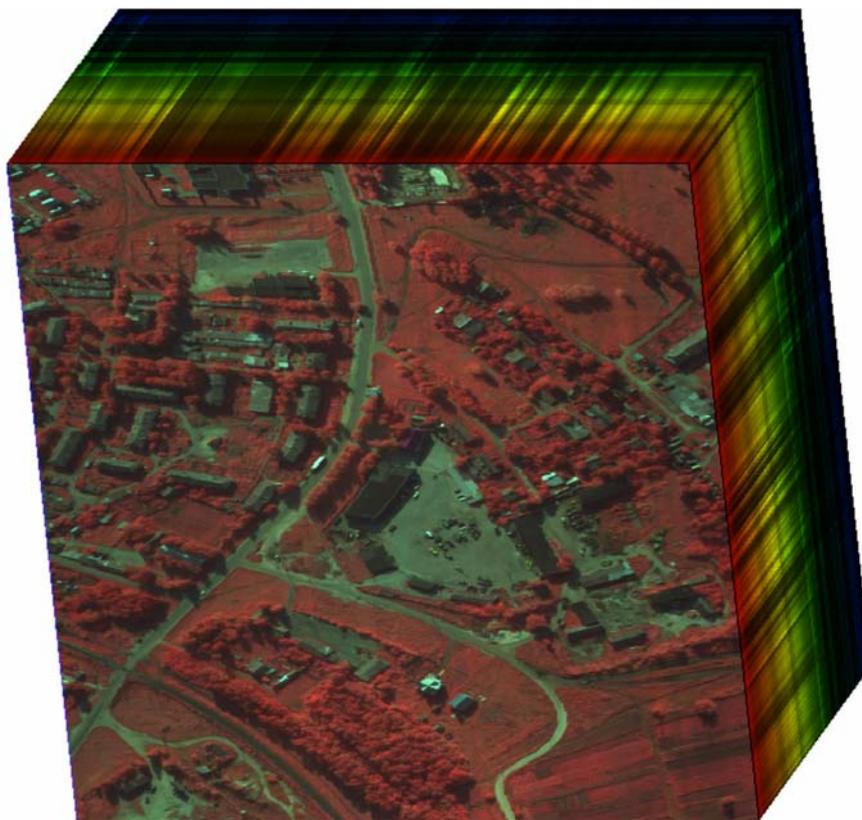


Рис. 1. Гиперкуб данных самолетной съемки Земли

Мера (1) характеризует среднеквадратическое отличие яркостей соседних элементов изображения с координатами (m, n) и $(m+1, n)$. Мера (2) описывается коэффициентом корреляции между спектральными характеристиками, определенными для тех же соседних элементов. Мера (3) численно равна углу в K -мерном спектральном пространстве.

Покажем, что мера (1) путем линейных преобразований спектральных характеристик может быть приведена к (2). Для этого рассмотрим две точки гиперспектрального снимка B_{mnk} и $B_{m+1,nk}$ и приведем соответствующие им частотные характеристики к единым среднему \bar{B} и СКО σ , найденным по всему гиперкубу. Выполним линейные преобразования B_{mnk} и $B_{m+1,nk}$ в B_{mnk}^* и $B_{m+1,nk}^*$:

$$B_{mnk}^* = a_{mn} B_{mnk} + c_{mn},$$

$$B_{m+1,nk}^* = a_{m+1,n} B_{m+1,nk} + c_{m+1,n},$$

где a и c – коэффициенты. Запишем выражения для средних значений: $B_{mn}^* = \bar{B} = a_{mn} \bar{B}_{mn} + c_{mn}$, $B_{m+1,n}^* = \bar{B} = a_{m+1,n} \bar{B}_{m+1,n} + c_{m+1,n}$ и СКО: $\sigma_{mn} = \sigma = a_{mn} \sigma_{mn}$, $\sigma_{m+1,n} = \sigma = a_{m+1,n} \sigma_{m+1,n}$. С учетом этих соотношений:

$$a_{mn} = \frac{\sigma}{\sigma_{mn}}, \quad a_{m+1,n} = \frac{\sigma}{\sigma_{m+1,n}},$$

$$c_{mn} = \bar{B} - \frac{\sigma}{\sigma_{mn}} \bar{B}_{mn}, \quad c_{m+1,n} = \bar{B} - \frac{\sigma}{\sigma_{m+1,n}} \bar{B}_{m+1,n}.$$

Выполним с учетом (5) и (6) преобразование (4) и на основе полученных B_{mnk}^* и $B_{m+1,nk}^*$ вычислим меру:

$$\Delta_{mn}^* = \sqrt{(B_{mn}^* - B_{m+1,n}^*)^2},$$

$$\text{где } B_{mnk}^* = \frac{\sigma}{\sigma_{mn}} (B_{mnk} - \bar{B}_{mn}) + \bar{B},$$

$$B_{m+1,nk}^* = \frac{\sigma}{\sigma_{m+1,n}} (B_{m+1,nk} - \bar{B}_{m+1,n}) + \bar{B}. \quad (8)$$

После преобразований получим

$$(B_{mnk}^* - B_{m+1,nk}^*)^2 = \sigma^2 \cdot \left[\left(\frac{B_{mnk} - \bar{B}_{mn}}{\sigma_{mn}} \right)^2 - 2 \frac{(B_{mnk} - \bar{B}_{mn})(B_{m+1,nk} - \bar{B}_{m+1,n})}{\sigma_{mn} \sigma_{m+1,n}} + \left(\frac{B_{m+1,nk} - \bar{B}_{m+1,n}}{\sigma_{m+1,n}} \right)^2 \right].$$

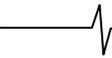
Можно показать, что усреднение последнего выражения по $k = 1, K$ приводит к формуле:

$$\Delta_{mn}^* = \sigma \sqrt{2(1 - \rho_{mn})}. \quad (9)$$

Выражение (9) устанавливает связь между двумя мерами близости элементов изображения Δ_{mn}^* , вычисляемой по формуле (7), и ρ_{mn} , определяемой (2). С учетом этого для дальнейшего анализа будем использовать меру (2).

Для определения меры близости элемента ГСИ B_{kmn} с тем или иным набором окружающих элементов можно воспользоваться операторами, используемыми для анализа в пространственной области: операторы градиента, Лапласа, Собела, Кирша и др. С привлечением натурной ГСИ выполнено сопоставление частотно-пространственных мер (2) и (3) в форме оператора градиента, как между собой, так и с оператором градиента, представленном в традиционном виде – в пространственной области. В этом случае меры (2), (3) в форме оператора градиента представляются как

$$\tilde{\rho}_{mn} = \sqrt{\rho_{mn}^2 + \rho_{m,n+1}^2}, \quad \tilde{\alpha}_{mn} = \sqrt{\alpha_{mn}^2 + \alpha_{m,n+1}^2}, \quad (10)$$



а собственно оператор градиента принимает значение

$$G_{mn} = \sqrt{(\bar{B}_{mn} - \bar{B}_{m+1,n})^2 + (\bar{B}_{mn} - \bar{B}_{m,n+1})^2}. \quad (11)$$

На рис. 2 в качестве примера в виде изображений представлены меры сходства точек ГСИ: слева – для

оператора градиента (11), справа – для частотно-пространственного корреляционного оператора (2). Как видим, оператор (2) дает лучшие результаты и может быть эффективно использован при реализации различных подходов сегментации гиперспектральных изображений.

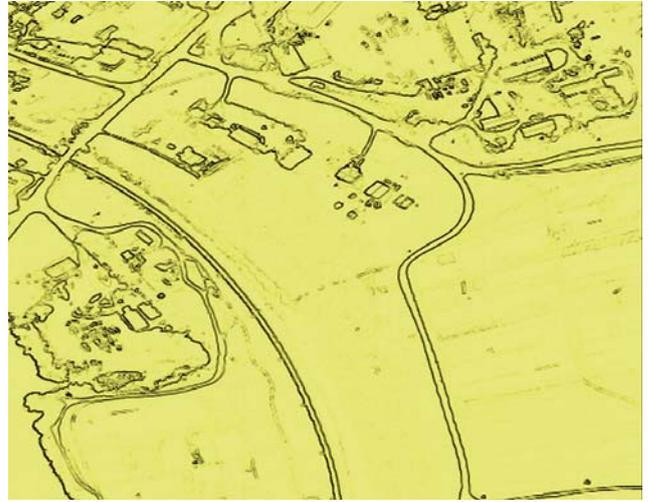
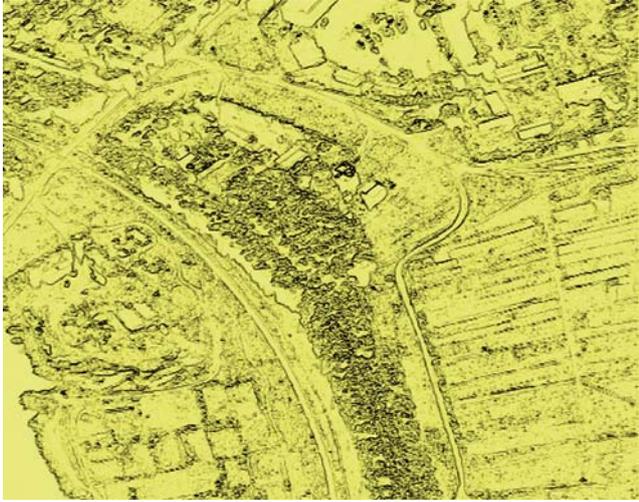


Рис. 2. Графическое отображение мер сходства (11) и (2)

Комплексирование ГСИ

При гиперспектральной съемке объекты наблюдаемой сцены совершенно по-разному отображаются в различных спектральных диапазонах. Часто наблюдается, что некоторый объект отображается более светлым по отношению к окружающей сцене в одних диапазонах и более темным в других. В результате этого при панхроматической съемке, при которой осуществляется интегрирование лучистой энергии в широком диапазоне спектра, такие объекты значительно теряют контраст либо вовсе становятся незаметными на снимках. Поэтому, в отличие от панхроматических снимков, ГСИ дают возможность сформировать некое новое изображение, в котором с высокой четкостью отображаются все объекты наблюдаемой сцены. Этот факт можно объяснить и аналитически, сопоставляя некоторые меры различимости объектов при панхроматической и гиперспектральной съемке.

Рассмотрим два элемента ГСИ B_{mnk} и B_{stk} , $k = \overline{1, K}$. В качестве меры различимости этих элементов на панхроматическом снимке естественно принять

$$d_{\Pi} = \left| \frac{1}{K} \sum_{k=1}^K B_{mnk} - \frac{1}{K} \sum_{k=1}^K B_{stk} \right| = \left| \frac{1}{K} \sum_{k=1}^K (B_{mnk} - B_{stk}) \right|. \quad (12)$$

Для ГСИ в качестве меры различимости может быть принято расстояние между рассматриваемым точками в K -мерном спектральном пространстве:

$$d_{\Gamma} = \sqrt{\frac{1}{K} \sum_{k=1}^K (B_{mnk} - B_{stk})^2}. \quad (13)$$

Эффективность использования ГСИ для комплексирования охарактеризуем отношением

$$\gamma = \frac{d_{\Pi}}{d_{\Gamma}} = \frac{\frac{1}{K} \sum_{k=1}^K |B_{mnk} - B_{stk}|}{\sqrt{\frac{1}{K} \sum_{k=1}^K (B_{mnk} - B_{stk})^2}} = \sqrt{\frac{(\overline{|B_{mn} - B_{st}|})^2}{K(B_{mn} - B_{st})^2}}. \quad (14)$$

Всегда $\gamma \geq 1$. Лишь при $B_{mnk} - B_{stk} = const$ для всех значений $k = \overline{1, K}$: $\gamma = 1$, и гиперспектральная съемка не дает дополнительной информации для повышения четкости изображения. По мере взаимной компенсации приращений спектральных компонент d_{Π} стремится к нулю, и создаются наилучшие условия для комплексирования гиперспектральных данных. В этом предельном случае

$$\gamma = \frac{1}{\sqrt{K}} \sqrt{1 + \frac{2 \sum_{j=1}^K \sum_{i=j+1}^K (B_{mnj} - B_{stj})(B_{mni} - B_{sti})}{\sum_{k=1}^K (B_{mnk} - B_{stk})^2}} = \frac{1}{\sqrt{K}}. \quad (15)$$

Рассмотрим один из апробированных алгоритмов комплексирования спектральных изображений [1], обобщенный на случай решения этой задачи с использованием ГСИ. Разделим изображения гиперкуба $B \in \{B_k, k = \overline{1, K}\}$ на низко- и высокочастотные части: $B_k = B_{kH} + B_{kB}$, $k = \overline{1, K}$. Для этого путем расфокусировки B_k , $k = \overline{1, K}$, сформируем низкочастотные составляющие B_{kH} , $k = \overline{1, K}$, а затем – высокочастотные составляющие $B_{kB} = B_k - B_{kH}$, $k = \overline{1, K}$. После этого путем взвешенного суммирования скомплексирем низкочастотные составляющие в виде изображения D_H :

$$D_H = \sum_{k=1}^K \alpha_k B_{kH}, \quad (16)$$

где коэффициенты α_k вычисляются для каждого элемента изображения D_H и характеризуют уникальность исходных пикселей:

$$\alpha_k = \frac{\sum_{j=1}^K \delta_{kj}}{\sum_{k=1}^K \sum_{j=1}^K \delta_{kj}}, \quad (17)$$

где $\delta_{kj} = |B_{kH} - B_{jH}|$, т.е. чем больше яркость элемента одного изображения отличается от яркостей этого же элемента в других изображениях, тем с большим весом пиксель участвует в формировании элемента результирующего изображения. В результате формируется низкочастотная компонента D_H результирующего изображения D . Формирование высокочастотной компоненты D_B результирующего изображения осуществляется следующим образом:

$$D_B = s \cdot \max_k |B_{kB}|, \quad k = \overline{1, K},$$

где s – знак выражения $\sum_{k=1}^K \alpha_k B_{kB}$.

В итоге результат комплексирования ГСИ формируется как сумма скомплексированных низко- и высокочастотных составляющих: $D = D_B + D_H$. На рис. 3 представлены: слева – в псевдоцвете панхроматическое изображение, справа – результат комплексирования ГСИ.



Рис. 3. Результат комплексирования ГСИ

Комплексирование ГСИ с данными высокодетальной съемки

Обычно в процесс формирования ГСИ осуществляется синхронная панхроматическая и (или) спектрзональная съемка Земли с многократно большей пространственной разрешающей способностью (детальностью). Стоит задача комплексирования видеоинформации, имеющей высокое спектральное разрешение, с данными высокодетальной съемки Земли с целью получения ГСИ с высокими и спектральным, и пространственным разрешением.

Пусть $D \in \{D_i(r, s), i = \overline{1, I}, I < K\}$ – спектрзональное изображение с I каналами, определенное в узлах частой прямоугольной решетки (r, s) , $r = \overline{1, R}$, $s = \overline{1, S}$, а $B^* \in \{B_k, k = \overline{1, K}\}$ – гиперспектральное изображение с более низким пространственным разрешением, зарегистрированное в узлах редкой прямоугольной решетки (m, n) , $m = \overline{1, M}$, $n = \overline{1, N}$, $m = \mu r$, $n = \mu s$, где μ – целочисленный масштабный коэффициент ($\mu = 2, 3, 4, \dots$). Рассмотрим алгоритм комплексирования изображений B и D .

Алгоритм основан на оценке функции распределения спектральной яркости для любой точки земной поверхности $E(\lambda)$ и формировании на ее основе спектральных компонент $B_k, k = \overline{1, K}$:

$$B_k = \int_{\lambda \in \Lambda_k} E(\lambda) S(\lambda) d\lambda, \quad k = \overline{1, K}, \quad (18)$$

где $S(\lambda)$ – функция относительной спектральной чувствительности видеодатчика, измеряемая на стадии предполетных испытаний, Λ_k – k -й спектральный диапазон.

После оценки $E(\lambda)$ формируется ГСИ $B_c \in \{B_{ck}, k = \overline{1, K}\}$ с высоким пространственным разрешением путем модуляции спектрзональным изображением D спектральных компонент B_k :

$$B_{ck} = D_i \frac{\int_{\lambda \in \Lambda_k} E(\lambda) S(\lambda) d\lambda}{\int_{\lambda \in \Lambda_i} E(\lambda) S(\lambda) d\lambda}, \quad k = \overline{1, K}, \quad i = \text{ent}\left(k \frac{I}{K}\right), \quad (19)$$

где Λ_k и Λ_i – соответственно спектральные диапазоны изображений B_k и D_i .

Для экспериментальных исследований алгоритма выполнено моделирование ГСИ низкого пространственного разрешения B^* путем прореживания отсчетов исходного ГСИ (см. рис. 1) в 5, 10, 20 и 30 раз. Затем из B^* в частой решетке (r, s) формировался гиперкуб данных $\tilde{B} \in \{\tilde{B}_k, k = \overline{1, K}\}$ путем пространственной билинейной интерполяции. Многозональное изображение D формировалось в решетке (r, s) как результат усреднения по k в I поддиапазонах данных исходного гиперкуба B .

На рис. 4 представлен результат комплексирования ГСИ B^* и многозонального изображения D в виде изображения B_c .

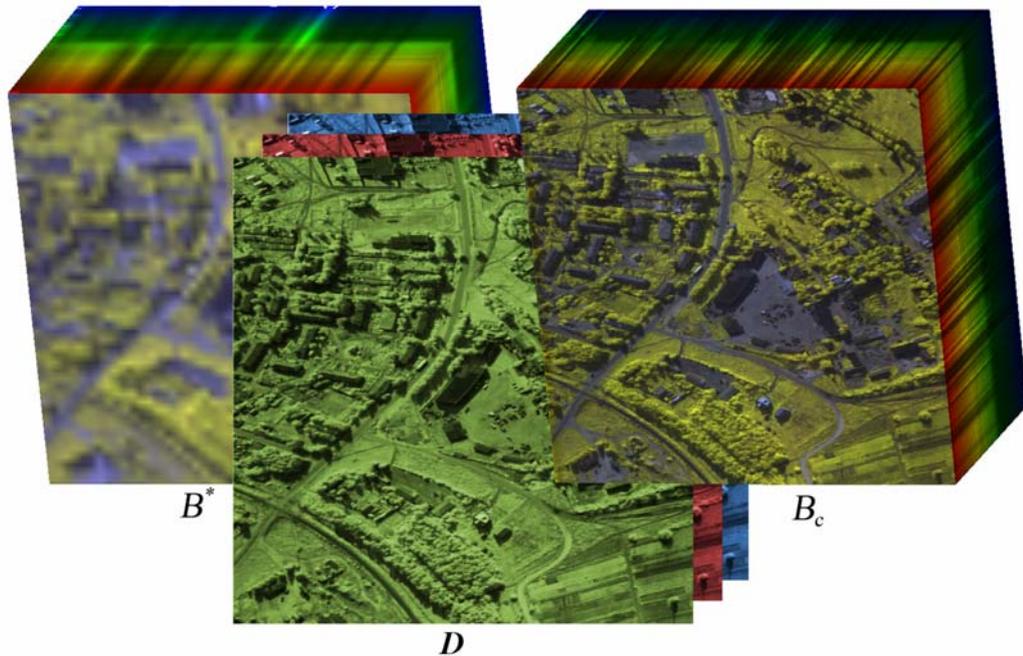


Рис. 4. Результат комплексирования B^* с D

Принципиально важным является вопрос о точности восстановления спектральных характеристик в точках (r, s) , не совпадающих с узлами решетки (m, n) . При известной функции спектральной чувствительности многозонального датчика $\alpha(\lambda)$:

$$\tilde{B}_k = \int_{\lambda \in \Lambda_k} \alpha(\lambda) D_i d\lambda, \quad k = \overline{1, K}, \quad i = \text{ent} \left(I \frac{k}{K} \right). \quad (20)$$

С другой стороны,

$$\bar{B}_{ck} = D_i, \quad k \in \left[(i-1) \frac{K}{I} + 1, i \frac{K}{I} \right], \quad (21)$$

где \bar{B}_{ck} – усреднение по k , принадлежащему указанному интервалу.

Число уравнений (20) равно MN , а число уравнений (21) равно RSI , всего уравнений $RSI+MN$, а число неизвестных KRS . Поэтому, чем больше привлекается спектрозональных снимков (параметр I), тем больше уравнений (20, 21) можно составить и тем точнее могут быть восстановлены спектральные характеристики в узлах (r, s) , не совпадающих с узлами (m, n) .

Выполнены экспериментальные исследования точности восстановления спектральных характеристик. Для этого во всех узлах частой решетки (r, s) сопоставлялись спектральные характеристики исходного гиперкуба (см. рис.1) и результата комплексирования в виде изображения B_c (см. рис.4). В качестве меры близости E использовалось среднее отношение модуля разности элементов полученного ГСИ B_c и исходного гиперкуба B , к уровню яркости элемента на изображении B :

$$E = \frac{1}{KRS} \sum_{r=1}^R \sum_{s=1}^S \sum_{k=1}^K |B_{krs} - B_{ckrs}| / B_{krs}.$$

При $K=105$, $\mu = 10$ получены следующие оценки:

- при комплексировании с однозональным (панхроматическим) снимком ($I = 1$): $E = 0.6624$ %;
- при комплексировании с 3-х зональным снимком ($I = 3$): $E = 0.4528$ %.

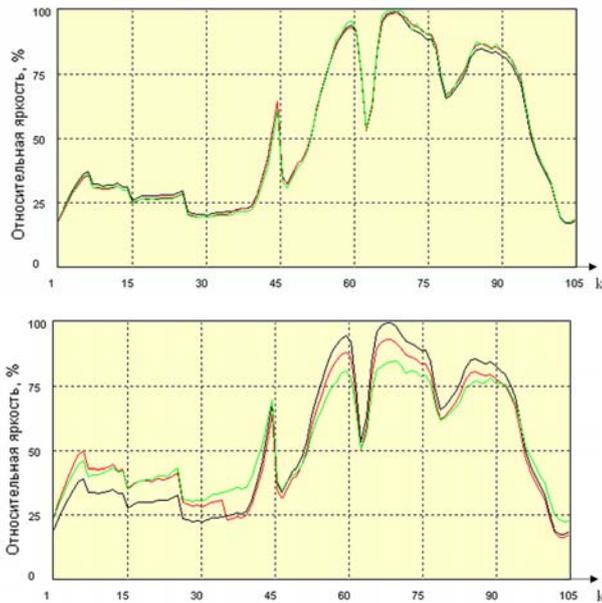
Гистограммы ошибки E приведены на рис. 5. Малая величина E обусловлена малым числом отсчетов с большими отклонениями от эталона, поскольку они наблюдаются только на границах объектов, доля которых невелика.



Рис. 5. Гистограмма распределения ошибки E

Проведен анализ зависимости величины ошибки E от μ . При $\mu = 5$: $E = 0,68\%$; $\mu = 10$: $E = 0,91\%$; $\mu = 20$: $E = 1,16\%$; $\mu = 30$: $E = 1,27\%$.

Для более детального изучения механизма работы алгоритма были построены функции спектральной яркости конкретных точек эталонного изображения B и результатов обработки B_c (рис.6): вверху – в узловой точке частой решетки (r, s), внизу – в центральной точке, наиболее удаленной от узловых. Как и следовало ожидать, результаты комплексирования с 3-х зональным снимком ближе к эталонным спектральным характеристикам, чем результат обработки по однозональному (панхроматическому) изображению.



- Эталонная спектральная характеристика
- Комплексирование по 3-х каналному снимку
- Комплексирование по панхроматическому снимку

Рис. 6. Спектральные характеристики точек ГСИ

Выводы

1. Рассмотрены меры сходства элементов гиперспектральных изображений, которые в отличие от традиционных мер, используемых для спектральнозональных и панхроматических изображений, основаны на корреляционном и среднеквадратическом сопоставлении спектральных характеристик различных точек ГСИ.

Экспериментально показано, что использование ГСИ для сегментации объектов позволяет более качественно решить эту задачу.

2. Рассмотрен алгоритм комплексирования ГСИ с целью получения нового изображения, на котором с высокой четкостью отображаются все объекты наблюдаемой сцены. Алгоритм основан на разделении ГСИ на низко- и высокочастотные компоненты и их отдельном комплексировании с учетом степени уникальности точек изображения.

3. Рассмотрен алгоритм комплексирования ГСИ с низким пространственным разрешением с более детальными спектральнозональными изображениями. Предложена методика оценки точности решения этой задачи. Экспериментально получены зависимости точности восстановления спектральных характеристик от соотношения разрешающих способностей. Так, например, при пятикратном отличии разрешающих способностей ГСИ и трехзонального снимков средняя ошибка восстановления спектральной характеристики составила $E = 0,68\%$.

Работа проводилась в рамках реализации ФЦП «Научные и научно-педагогические кадры инновационной России» на 2009 – 2013 годы (госконтракт № П2422).

Литература

1. Злобин В.К., Еремеев В.В.. Обработка аэрокосмических изображений. – М.: Физматлит, 2006. 286 с.
2. Yuhas, R.H., Goetz, A. F. H., and Boardman, J. W., 1992, Discrimination among semiarid landscape endmembers using the spectral angle mapper (SAM) algorithm. In Summaries of the Third Annual JPL Airborne Geoscience Workshop, JPL Publication 92-14, vol. 1, pp. 147-149.

SPECIFICS OF ANALYSIS AND PROCESSING OF INFORMATION FROM SATELLITE HYPERSPECTRAL EARTH IMAGING SYSTEMS

A number of problems on analysis of hyperspectral images (HSI) of earth surface were examined: measurements of HSI elements similarities, complexing of HSI with images from different imaging systems, HSI object clusterization. Results of research of these algorithms were presented with airborne hyperspectral earth surface images involvement.

СЕМЕЙСТВО МНОГОФУНКЦИОНАЛЬНЫХ СИСТЕМ ОБРАБОТКИ ВИДЕОИЗОБРАЖЕНИЙ «ОХОТНИК»

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

Блохин А.Н., начальник лаборатории научно-конструкторского центра видеоконьютерных технологий (НКЦ ВКТ) Государственного Рязанского приборного завода (ГРПЗ), e-mail: hunter-rzn@yandex.ru

Костяшкин Л.Н., к.т.н. доцент, директор-главный конструктор НКЦ ВКТ ГРПЗ, e-mail: hunter-rzn@yandex.ru

Романов Ю.Н., заместитель директора НКЦ ВКТ ГРПЗ, e-mail: hunter-rzn@yandex.ru

Шапка С.В., начальник лаборатории НКЦ ВКТ ГРПЗ, e-mail: hunter-rzn@yandex.ru

Ключевые слова: обработка видеоизображений, автоматическое обнаружение, захват и сопровождение, оптико-электронные системы, семейство «Охотник».

Введение

Комплексный подход при проектировании и разработке обзорных и обзорно-прицельных оптико-электронных систем (ОЭС) нового поколения для разнообразных систем вооружений и военной техники (ВВТ) связан как с новыми решениями общесистемного характера, так и с решениями по основным компонентам ОЭС. Интеллектуальным ядром многих из них являются системы обработки видеoinформации в реальном времени (СОВИ), разработка и серийное производство которых ведутся на ФГУП «Государственный Рязанский приборный завод». При их проектировании и разработке должен быть решен ряд проблем, определяющих в конечном счете эффективность ОЭС в целом:

- обоснование требований к выбору системы видеодатчиков, входящих в состав ОЭС, с учетом ее назначения;
- определение набора реализуемых функций по обработке изображений;
- выбор общесистемных принципов организации системы обработки видеоизображений (СОВИ): архитектуры, вычислительной платформы, элементной базы, интерфейсов и др.;
- разработка алгоритмического и программного обеспечения СОВИ;
- выбор конструктивного исполнения СОВИ в соответствии с предъявляемыми к ней требованиями.

Решения перечисленных проблем многовариантны и взаимозависимы, их приемлемость определяется соответствием заданным требованиям применения. Сама направленность функциональных задач СОВИ в обзорных и обзорно-прицельных ОЭС определяется их назначением, а также ориентированностью – для работы оператора или для работы в автоматическом режиме, что влияет на состав математических операторов обра-

Рассмотрено семейство многофункциональных систем обработки видеоизображений «Охотник», разработанное на ФГУП «ГРПЗ». Охарактеризованы выполняемые изделиями семейства функции по обработке телевизионных и тепловизионных изображений, такие как улучшение видения оператором, в том числе комплексированием их, автоматическое обнаружение и автоматическое сопровождение целей, координатные преобразования и др. Отражены особенности программно-алгоритмического обеспечения и вычислительных платформ, использованных в изделиях семейства.

ботки и требования к ним. Востребованной для большинства применений является многофункциональность обработки, определяемая целями и обстоятельствами конкретного использования СОВИ в ОЭС [1]. Обобщенный по ряду применений набор функций обработки изображений включает в себя:

- **улучшение видения** монохромных и цветных телевизионных (ТВ) и тепловизионных (ТПВ) изображений;
- **комплексирование** разномасштабных телевизионных (ТВ) и тепловизионных (ТПВ) изображений, в перспективе включая и радиолокационные с выходом на совмещение комплексированного изображения с 3-мерной картой местности и в пределе формированием синтетического изображения (технология ESVS – enhanced synthetic vision system);
- **координатные преобразования** изображений, включающие **электронную стабилизацию** картины поля зрения видеодатчика, **компенсацию поворота** изображения, **электронное масштабирование** изображения, **электронное выравнивание** полей зрения и **юстировку** линий визирования;
- визуальное или автоматизированное **обнаружение, захват и автоматическое сопровождение** целей (с управлением линией визирования видеодатчиков).

При общности основных функциональных задач обработки в различных ОЭС получение единого технического решения в виде универсального ядра для разных применений ограничивается спецификой и особенностями фоноцелевой картины конкретных областей применения, наличием разных по составу и параметрам видеодатчиков, используемыми интерфейсами и логикой взаимодействия, числом видеовходов и видеовыходов и связанной с этим параллельностью поканальной

обработки, различными требованиями по условиям эксплуатации и конструкции в соответствии с группой исполнения, экономическими соображениями и др. Все это приводит к множественности реальных объектоориентированных изделий, которые объединены в семействе многофункциональных СОВИ «Охотник». В основу их построения положены общие системотехнические принципы и максимальная унификация аппаратных и программно-математических решений. К настоящему времени семейство насчитывает свыше полутора десятка типов изделий для систем прицеливания и управления ракетно-пушечным вооружением, систем наблюдения и контроля, для поисковых систем, систем пилотирования и вождения бронетехники различных объектов ВВТ, а также для гражданских применений.

Особенности базового программно-алгоритмического обеспечения

В состав практически всех изделий семейства «Охотник» входят в качестве базовых следующие функции: **улучшение видения (УВ)**, **автоматическое обнаружение (АО)** и **сопровождение целей (АС)**.

1. Алгоритмическое обеспечение функции УВ ориентировано на решение задач двух групп – преобразование исходных изображений с целью улучшения их восприятия при визуализации по отдельности для ТВ и ТПВ датчиков и комплексирование разноспектральных видеоизображений (технология Fusion). Цель данных преобразований – повышение информационных характеристик видеоизображений для усиления визуальных возможностей по предельным дальностям видения, повышения вероятности обнаружения и идентификации целей.

На процесс поиска и распознавания объектов и вероятность визуального восприятия оказывают влияние характеристики объекта (угловой размер, градиенты изменения яркости на краях крупных деталей изображения, расположение на экране, скорость движения, контраст относительно фона, характеристики наблюдателя: тренированность, утомление, возраст, интеллект, рабочая нагрузка, периферическая острота зрения, наличие предварительных инструкций, а также характеристики видеоинформационного тракта: пространственно-частотная передаточная характеристика, число строк, приходящихся на угловой размер объекта, количество воспроизводимых градаций). Зависимость вероятности визуального обнаружения от ряда факторов по результатам исследований [2] имеет вид

$$P = 1 - \exp\left(-\frac{CK^2\gamma^3L^{0.3}t}{(2\beta)^2}\right),$$

где C – некоторая константа, характеризующая конкретного наблюдателя; K – локальный контраст объекта; γ – угловой размер объекта; L – яркость фона; t – время поиска; 2β – угловое значение поля зрения. Исключив из рассмотрения отношение времени поиска на угловое значение поля зрения, вероятность обнаружения напрямую будет зависеть от значений локального контраста объекта, углового размера объекта и яркости фона изображения сцены.

Собственно математические операторы УВ, реализуемые в аппаратуре СОВИ, выполняют функции коррекции

передаточной функции тракта преобразования, изменяя его пространственно-частотную характеристику в желаемом направлении (Рис.1). При этом максимальный эффект УВ достигается с учетом индивидуальных особенностей видеодатчиков, тракта передачи видеоинформации и монитора, отображающего фоноцелевую картину.

Кроме перечисленных факторов, значительное влияние на характеристики обнаружения оказывают шумовые компоненты видеосигнала от датчиков (структурного или случайного шума) на всем изображении и в окрестности наблюдаемого объекта, а также не идеальности канала передачи данных. Это особенно проявляется при обнаружении и распознавании объектов на слабоконтрастных изображениях при низкой заметности целей в условиях плохой видимости или малых температурных контрастов, (туман, пыль, дымовые завесы, наличие осадков), при наличии пассивных и активных видов противодействия (маскировка, пиротехнические средства, прожектора и т.п.) [3].

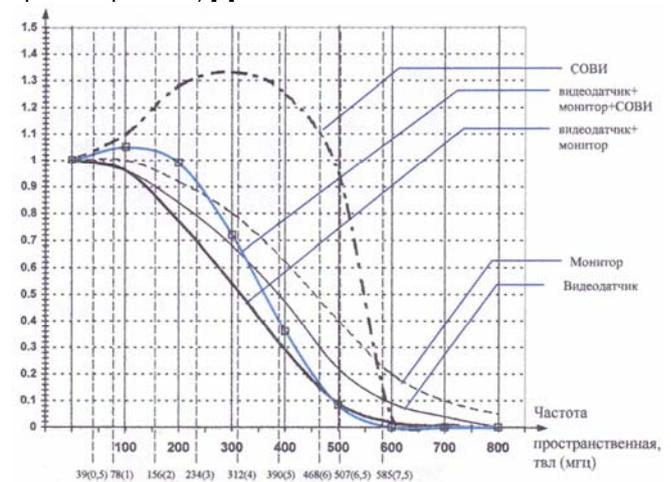


Рис.1 Пространственно-частотные характеристики отдельных компонент СОВИ

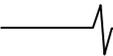
С учетом вышеизложенного, в состав алгоритмического обеспечения преобразований УВ первой группы включены операторы:

- адаптивное линейное яркостное преобразование, которое нормирует диапазон изменения выходного видеосигнала, увеличивает глобальный контраст и поддерживает максимально необходимое (256) или реально имеющееся количество информационных уровней выходного изображения;

- нелинейные яркостные преобразования, основанные на гистограммных преобразованиях (локальных и глобальных) с использованием различных законов, в том числе по специально разработанным шаблонам, позволяющим эффективно контрастировать изображения и нормализовывать их среднюю яркость;

- адаптивная временная фильтрация (рекурсивная фильтрация или накопление кадров с попиксельной межкадровой оценкой), обеспечивающая подавление случайных помех, особенно проявляющихся на уровне порогов чувствительности ТВ или ТПВ датчиков и повышающая соотношение сигнал/шум;

- пространственная фильтрация (ранговая или линейная), подавляющая структурные, импульсные и другие шумы видеодатчиков;



– дифференциальные преобразования яркостей, позволяющие усилить разборчивость детальной составляющей изображения.

Адаптивная комбинация различных преобразований, подбор их параметров под текущие условия наблюдения (в том числе автоматически), позволяют осуществить в конечном счете улучшение видения изображений от ТВ и ТПВ сенсоров за порогом чувствительности глаза.

Реализация алгоритмического обеспечения функции УВ в изделиях семейства «Охотник» при их использовании в ОЭС обеспечила:

- увеличение дальности обнаружения и распознавания малоконтрастных объектов/целей при ограниченно сложных метеоусловиях от 1.3 до 1.7 раз;
- снижение времени обнаружения-распознавания в 3 – 8 раз;
- расширение времени действия каналов технического зрения до 0.5 часа в весенне-летний период и до 1.0 часа в осенне-зимний период;
- снижение утомляемости оператора на 30-40%;
- снижение времени от обнаружения до момента применения оружия;
- повышение вероятности выполнения боевой задачи.

Приведенные оценки получены расчетным путем и подтверждены на практике в процессе испытаний и при эксплуатации ряда изделий семейства.

Эффективным средством повышения информативности визуализируемого изображения является комплексирование ТВ и ТПВ изображений (вторая группа алгоритмических преобразований - отдельный раздел УВ). Для этого необходимо точное совмещение полей зрения видеодатчиков. Изображения от них имеют характерные яркостно-геометрические особенности, обусловленные физической природой формирования изображений и характеристиками оптико-электронных приборов. ТВ-изображения (оптический диапазон) имеют естественный (привычный) вид для оператора, в них присутствуют линия горизонта, тени на объектах и всей фоноцелевой картине. В то же время, ТПВ-изображения (ИК-диапазон), не имеют теней, слабо различима линия горизонта, что затрудняет пространственно-объемное восприятие визуальной картины, но зато в них фиксируются характерные признаки объектов, отличающиеся тепловыми контрастами. Таким образом, комплексная обработка разноспектральной информации предоставляет дополнительные возможности при автоматизации обработки изображений и распознавания объектов. С учетом информативности каждого из видеодатчиков комплексирование может осуществляться несколькими вариантами [4].

Первый вариант - комплексирование (преобразование по схеме "полутонные изображения" - "информативные признаки") предполагает использование определенных признаков в каждом канале. Здесь могут не предъявляться жесткие требования к видеодатчикам как по вертикальному и горизонтальному разрешению, так и по количеству уровней и геометрическому представлению пикселя.

Второй вариант - комплексирование (преобразование по схеме "полутонные изображения" - "графиче-

ское изображение") использует совместное преобразование видеoinформации от разных датчиков. Этот уровень требует большего подобия датчиков по их характеристикам и условиям работы, а также одновременного восприятия фоноцелевой картины и синхронной выдачи информации.

Третий вариант - комплексирование предполагает получение результирующего изображения от различных сенсоров попиксельно по схеме "полутонное изображение" - "полутонное изображение". Третий вариант возможен при наличии датчиков с близкими по пространственному и контрастному разрешению параметрами, а результат комплексирования в этом случае представляется в виде объемного пространственного изображения, приближающегося к естественному. Этот вариант сложен как по алгоритмическому обеспечению, так и по аппаратным средствам реализации потоковой цифровой обработки в реальном времени.

Реализованный в изделиях семейства «Охотник» вариант комплексирования использует текущие характеристики ТПВ-изображения для дополнительного изменения текущих значений яркости телевизионной картинки. ТПВ-датчик и выполняет функцию обнаружителя потенциальных целей, препятствий, характерных элементов фоноцелевой картины. Такой подход позволяет получить естественное полутонное комплексированное изображение с детальной информацией от ТВ-датчика и повышенным локальным контрастом, а его алгоритмическое обеспечение обеспечивает последующую аппаратную реализацию в реальном времени с максимально возможными характеристиками (768x576 элементов разложения при 256-ти градациях уровня серого). На комплексированном видеоизображении обеспечивается проработка текстурной составляющей подстилающей поверхности за счет внесения информации о температурных контрастах, сохраняется естественное объемное восприятие, присущее телевизионному изображению за счет наличия теней. Возможен также эффект stereopsis, при разнесении ТВ и ТПВ датчиков, который улучшает дополнительно пространственное и объемное восприятие на дальностях до 30 - 50 метров, что особенно важно для обзорных систем управления подвижными объектами или летательных аппаратов при маловысотном полете и посадке на необорудованные площадки. В аппаратуре осуществляется попиксельное совмещение изображений за счет выравнивания полей зрения и юстировки линий визирования, синхронизация видеопотоков данных от ТВ и ТПВ датчиков. Предварительно осуществляется поканальная обработка изображений для поддержания постоянного контраста, стабилизации средней яркости, повышения четкости и снижения уровня шума в выводимом изображении при сохранении максимального количества градаций уровней серого [5].

На восприятие изображений (как исходных, так и преобразованных с целью улучшения видения) в целом оказывают отрицательное влияние и такие факторы как: отсутствие или недостаточность стабилизации изображения, малые угловые размеры изображения интересующего объекта, неудобное для работы с ним угловое положение изображения, нелинейные геометрические искажения и др. В целях коррекции визуализируемых

изображений в состав алгоритмического обеспечения изделий семейства «Охотник» входит раздел координатных преобразований, включающих электронную стабилизацию изображений, компенсацию углового поворота, электронное масштабирование, электронную юстировку линий визирования сенсоров, выравнивание полей зрения и др.

Алгоритмическое обеспечение предусматривает ручной и автоматический способы настройки аппаратных средств задачи УВ на максимально высокое качество выходного изображения.

При ручном способе настройки возможен выбор преобразований: яркостная коррекция – парирование «белого», парирование «черного»; временная фильтрация – изменение глубины рекурсии. Виды преобразований, апертуры, коэффициенты выбираются в соответствии с решаемой задачей, типом датчика (ТВ, ТПВ) и условиями наблюдения.

При автоматической адаптации алгоритмической системы используются более сложные решения, обеспечивающие «интеллектуальную» настройку аппаратуры на максимально информативное для оператора выходное изображение по нормированному контрасту, средней яркости при максимально возможном значении сигнал/шум.

2. Алгоритмическое обеспечение задач обнаружения, захвата и автоматического сопровождения целей практически во всех изделиях семейства «Охотник» представляет собой основной раздел общего алгоритмического обеспечения каждого из них. Решение задачи автоматического обнаружения цели сокращает время реакции оператора на анализ фоноцелевой обстановки. Автоматическое сопровождение цели в замкнутом контуре значительно повышает точность работы ОЭС при работе ракетно-пушечным вооружением.

В первых изделиях семейства (СОИ-ОПК для вертолетов типа Ми-8МТКО) круг решаемых задач ограничивался сопровождением наземных неподвижных площадных целей на неравномерном фоне. Задача решалась с помощью корреляционного разностного алгоритма, описанного в [6,7], на базе вычисления разностной критериальной функции отличия текущего изображения объекта от эталонного

$$F(\alpha, \beta) = \sum_{i, j \in H} |h(i, j) - l(i + \alpha, j + \beta)|,$$

где $l(i, j)$ – текущее изображение, $h(i + \alpha, j + \beta)$ – эталонное изображение цели, α, β – параметры, характеризующие смещение центра объекта относительно начала координат. Зона поиска изображения цели ограничена, частота обработки информации – 25 Гц. Для повышения устойчивости работы алгоритма при изменениях ракурса цели, движении носителя в алгоритме использовано межкадровое сглаживание (фильтрация) эталонного изображения цели, позволяющее учесть изменения цели, накапливающиеся в процессе ее автосопровождения.

В усовершенствованном варианте этого изделия (СОВИ ОПС), наряду с корреляционным алгоритмом, введен алгоритм сопровождения, использующий статистическую сегментацию обрабатываемого изображения [7]. Алгоритм предназначен для слежения за малораз-

мерными протяженными целями на равномерном фоне (небо) и за протяженными контрастными целями на неравномерном фоне (земля, море). Его основу представляет байесовский критерий принятия решений о принадлежности точки на изображении фону или цели. На изображении выделяются две области: область окна и область рамки так, чтобы объект полностью помещался в окне, а в рамке находились только точки фона. Классификация точек проводится в соответствии с выражением

$$\hat{P}(O) = \alpha \cdot A \cdot \hat{P}(P),$$

где $P(O)$ – оценка вероятности принадлежности точки к цели, $P(P)$ – оценка вероятности принадлежности точки к фону, α – априорная вероятность принадлежности точки к цели, A – стоимость ошибочной классификации.

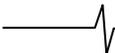
В процессе работы алгоритма важно поддержание соответствия между площадью окна и площадью рамки при построении гистограмм, обеспечиваемого введенной процедурой адаптивной подстройки размеров окна и рамки. Для захвата цели на сопровождение в начальный момент времени оператор должен «стробировать» цель (наложить на нее прицельный маркер определенного размера). При этом обеспечивается устойчивое слежение за целью. Чтобы облегчить и ускорить процедуру захвата в алгоритм введен циклический перебор размеров прицельного маркера во время захвата. Этим удастся избавить оператора от необходимости подбора размеров прицельного маркера вручную.

В корреляционном алгоритме слежения и в алгоритме, основанном на сегментации, имеется возможность продолжения слежения за целью при потере оптической связи с ней. Для прогнозирования положения цели используются оценки траекторного фильтра, формируемые в процессе слежения. При работе в замкнутом контуре сопровождения прогнозируемое положение поворотной платформы формируется по данным о её скорости и ускорении, накопленным к моменту пропадания цели. Допускается прогнозирование положения цели в течение 5-6 секунд. Увеличение интервала времени нецелесообразно в связи с возрастанием ошибок в прогнозе из-за различий в скоростях перемещения цели и носителя.

По истечении времени прогноза выполняется поиск цели в расширенной зоне анализа. Если цель найдена, слежение продолжается, в противном случае слежение завершается.

Дальнейшее совершенствование алгоритмического обеспечения связано с увеличением максимального размера обрабатываемого эталонного изображения, увеличением зоны поиска и обработкой одновременно нескольких эталонных изображений.

В изделии АТТ (применение на вертолете Ми-28Н и самолете Су-34) существенно развит корреляционный алгоритм в части одновременной обработки до 25 эталонных изображений (точек привязки), что позволило реализовать режим «электронной» стабилизации изображения. Включение режима «электронной» стабилизации увеличивает дальность обнаружения и распознавания типовых целей от 2 до 4 км в зависимости от сложности фоноцелевой картины. Эффективность вве-



денных алгоритмических усовершенствований подтверждена реальной работой свыше сотни произведенных изделий на реальных объектах ВВТ.

Важным фактором в работе любой СОВИ является минимизация участия оператора в управлении ею. Поэтому в изделии, применяющемся на вертолете Ка-52, реализован алгоритм автоматического обнаружения движущихся целей [8], позволяющий значительно упростить действия оператора при захвате цели на сопровождение.

Ядром алгоритма обнаружения является вычисление оценок фонового изображения и формирование на их основе бинарной картинке с выделенным движущимся объектом. Алгоритм включает в себя этапы межкадрового оценивания фонового изображения, вычисление вероятностей значимых изменений, пороговую обработку, специальные процедуры разметки бинарных изображений и оценки параметров сегментов. В результате выделяются движущиеся цели с наложенными прицельными маркерами. Оператору необходимо лишь выбрать нужную цель и начать автосопровождение.

Если на изображении нет движущихся целей, то используется алгоритм обнаружения неподвижных целей, задача которого заключается в выделении на изображении областей, однородных по яркости и стробирование их прицельной маской.

Важным направлением разработки нескольких изделий семейства «Охотник» оказалось введение в алгоритмическое обеспечение возможности работы по точечным малококонтрастным воздушным целям. Эта функция реализована в изделиях ОВИ-М2 (применение в ЗРК «Квадрат»), ОВИ-М3 (применение в ЗРК «Бук-М2Э»). В алгоритмах обнаружения точечных целей используется параллельная фильтрация текущего изображения масками, учитывающими наличие объекта на изображении и его отсутствие [9]. Входное изображение $I(i, j)$ обрабатывается линейным фильтром, имеющим маску w_1 размерности $q_{11} \times q_{12}$. Параллельно с этим оно сглаживается фильтром с маской w_2 размерностью $q_{21} \times q_{22}$, причем $q_{21} > q_{11}$, $q_{22} > q_{12}$. По специальному решающему правилу точки разделяются на точки фона и цели. Алгоритм позволяет обнаруживать цели с размерами от 2×1 до 32×32 элементов разложения. Это позволяет в некоторых случаях использовать параллельно данный алгоритм и байесовский алгоритм сегментации. Совместное применение алгоритмов значительно повышает устойчивость слежения за целями.

Применение разработанного алгоритмического аппарата позволяет изделиям семейства «Охотник» эффективно и с высокой устойчивостью решать поставленные задачи и обеспечивает следующие основные технические характеристики и возможности:

- автоматическое обнаружение от 1 до 8 воздушных и движущихся целей;
- автоматизированный захват целей на сопровождение с адаптацией размера следящего строба;
- анализ возможности захвата целей на сопровождение;
- захват цели на автосопровождение размерами от 2×2 до 256×256 элементов разложения (э.р.);
- слежение за целью при изменении её линейных

размеров от 2×2 до 256×256 э.р.;

- слежение за целью при скорости ее движения до 250 э.р. в секунду;
- слежение за целью при изменении ее ракурса со скоростью до 30% в секунду;
- слежение за целью при изменении ее средней яркости со скоростью до 20 уровней яркости в секунду;
- поиск цели в расширенном следящем стробе при потере оптической связи с целью на время до 6 с;
- слежение за целями при отношении сигнал/шум не менее 3;

Практика использования аппаратуры семейства «Охотник» в различных конкретных применениях приводит к необходимости решения ряда проблем их комплексного взаимодействия в составе бортового оборудования объекта ВВТ, что должно обеспечиваться соответствующими индивидуальными расширениями программно-алгоритмического обеспечения отдельных изделий семейства.

Особенности вычислительной платформы семейства «Охотник»

Сложность математического обеспечения и необходимость вычислений в реальном времени определяют жесткие требования к вычислительной платформе в отношении производительности вычислений и, соответственно, к архитектуре, техническим решениям и элементной базе.

Вычислительная платформа системы обработки видеоинформации реализует алгоритмическое обеспечение всех функциональных задач в реальном времени с эквивалентной производительностью вычислений порядка 10-12 тысяч MIPS. Она представляет собой открытую систему и содержит ресурсы и возможности, необходимые для функциональной адаптации под условия конкретного применения. При организации вычислительной платформы доминирующим является принцип аппаратной реализации математического обеспечения функциональных задач цифровой обработки изображений, реализованный на базе реконфигурируемых процессоров (ПЛИС). Программные средства (цифровые процессоры обработки сигналов) используются здесь для реализации функций управления, обмена информацией и т.п., где присутствует сложная логическая разветвленность алгоритмов и не требуется высокая производительность вычислений. Это обеспечивает оперативную загрузку различных алгоритмов УВ и комплексирования разноспектральных изображений, координатных преобразований, обнаружения, выделения и измерения текущих координат целей, а также алгоритмов управления, как непосредственно в процессе работы системы, так и при необходимости изменения ее функциональных возможностей.

Алгоритмы преобразования изображений (УВ) для представления оператору, а также алгоритмы АО и АС целей, предъявляют ряд требований к реализующей их вычислительной платформе. На начальном этапе применения цифровой обработки изображений в оптических прицельных станциях использовались корреляторы с размером эталона не более чем 8×8 точек, работающие по зоне поиска 16×16 э.р. Далее требования увеличилось до

32x32 с зоной поиска 64x64. В настоящее время требуется реализовывать эталоны до 64x64 точки с зоной поиска 128x128 э.р., кроме того, появились требования по работе с датчиками цветных телевизионных изображений, что привело к увеличению потока обрабатываемого изображения более чем в два раза. Количество таких одновременно работающих корреляторов возрастает с 1-го до 20 и более. Подобный рост обусловлен как повышением требований к устойчивости работы следящих алгоритмов, так и изменением технических характеристик систем технического зрения, в первую очередь в части возрастания пространственного разрешения и кадровой частоты. Свою специфику вносит также работа в узких полях зрения на больших скоростях, особенно это проявляется в летательных аппаратах (из-за вибрации линии визирования появляется необходимость включения электронной стабилизации). Примерно так же растут требования ко всем видам цифровой обработки изображений.

Реализация алгоритмов в реальном времени для различных применений возможна несколькими различными вариантами.

Первый вариант – на основе процессорных вычислений при помощи DSP, имеющих в своем составе узлы для потоковой обработки. Это возможно, например, в процессорах Texas Instruments и Analog Devices, использующих VLIW-архитектуру (до 8 операционных узлов, способных работать параллельно, выполняя одну операцию за такт на частотах от 500 МГц до 1 ГГц). Однако структура операционных узлов данных процессоров ориентирована на определенный круг задач, на котором их эффективность максимальна. Любой “шаг в сторону” приводит как к простаиванию отдельных узлов процессорного конвейера, так и к необходимости производить вспомогательные действия, связанные с подготовкой операндов к вычислениям. В этом случае для реализации задач в полном объеме необходимо количественное наращивание операционной емкости, что приводит к неприемлемому росту габаритов, энергопотребления и снижению наработки на отказ.

Второй вариант – создание аппаратных средств, специализированных на конкретную задачу (ASIC для коррелятора, ASIC для обнаружителя и др.). Такой путь дает самые малогабаритные и малоэнергопотребляющие технические решения. Однако время на разработку специализированных СБИС на настоящий момент очень велико. Более того, как показывает практика, на этапе разработки (вплоть до запуска изделий в серию, а иногда и после, в процессе эксплуатации), возникает необходимость внесения корректив в действующие алгоритмы, что обусловлено рядом обстоятельств конкретного применения. На базе СБИС такая гибкость невозможна.

Третий вариант – программно-аппаратная платформа, сочетающая в себе гибкость перепрограммирования с необходимой производительностью специализированных микросхем. Такую возможность предоставляют программируемые интегральные логические схемы (ПЛИС). В сочетании с сигнальными процессорами, применяемыми в данном случае для решения задач с меньшим объемом данных (управление в контуре, построение и поддержание связанных списков обнаруживаемых в каждом кадре объектов и др.), удается достичь компромис-

са между гибкостью и производительностью платформы обработки. Жесткая схема аппаратных вычислений, необходимая для достижения высокого быстродействия, совмещается с гибкой и оперативной реконфигурацией вычислительной среды в рамках возможностей ПЛИС по технологии FPGA (field programmable gate array), обеспечивая возможность быстрой адаптации изделия к различным условиям его применения в составе обзорно-прицельных систем. Ряд внешних воздействующих факторов, способных нарушить функционирование устройства, требует применять технические решения, снижающие вероятность отказов устройства в целом до требуемой величины. Использование ПЛИС позволяет, с одной стороны, решать ряд последовательных задач на одной базе путем перезагрузки содержимого (без потери потоковой производительности), с другой стороны, возможно организовать параллельные структуры, необходимые для 2-х или 3-х кратного резервирования.

В качестве аппаратных элементов для потоковой видеообработки хорошо зарекомендовали себя ПЛИС производства фирмы Xilinx. На их базе реализованы все разработанные и поставляемые изделия семейства «Охотник». В качестве сигнальных процессоров применяются DSP фирм Texas Instruments и Analog Devices. Полученные результаты полностью удовлетворяют требованиям функциональных задач ЦОИ (технических заданий) для данных применений, с гарантийным запасом операционных ресурсов в случае усложнения задачи или при необходимости адаптации с целью применения на других изделиях.

Необходимо отметить имеющуюся необходимость использования высокоскоростных цифровых интерфейсов межсистемного (бортового) обмена. Она возникает как из-за невозможности применения старых аналоговых линий связи для передачи видеосигнала высокой четкости, так и из-за появляющейся в модернизируемых комплексах необходимости передачи по одной линии связи данных от нескольких источников видеосигнала. Использование ПЛИС позволяет легко задать любой протокол передачи данных (основанный на каком-либо стандарте, либо специфичный для конкретного применения, например, в ряде моделей семейства обеспечивается передача и прием видеосигналов по скоростному цифровому интерфейсу спецификации Fibre Channel Physical Standard (ANSI X3.230-1994).

Параллельно с наращиванием производительности вычислений острой является задача существенного сокращения массо-габаритных показателей изделий (например, для малоразмерных беспилотных летательных аппаратов). В этом случае эффективность использования операционных ресурсов становится особенно актуальной и реализация на одних только DSP исключается из-за вышеуказанного неэффективного использования схем процессоров при обработке сигнала. Заказные СБИС имеют большую операционную емкость, чем ПЛИС, однако последние дают возможность осуществлять быструю перезагрузку в процессе работы, позволяя тем самым «экономить» на задачах, не выполняющихся параллельно (разные операторы УВ, различные алгоритмы АО и АС целей). Современные ПЛИС позволяют выполнять частичную перезагрузку прошив-

ки, если недопустимо кратковременное отключение в процессе работы в части решения какой-либо функциональной задачи.

Разработанные изделия семейства «Охотник» с разнообразными наборами функций и возможностей для различных применений приведены в таблице 1.

Таблица 1.

Характеристики функциональности изделий семейства СОВИ «Охотник»

Тип изделия	Системы применения	УВ	К	ЭС	КК	ЭМ	АО	АЗ	ЭСл	АСЗК	Ц	Интерфейсы *	Группа исполнения	Литера
СОИ-ОПК	обзорная	+	-	-	-	-	-	-	+	+	1	RS-232	3.1.1	КД ГК
СОВИ-ОПС	обзорная	+	-	-	-	-	-	-	+	+	1	RS-232	3.1.1	КД ГК
СОИ-ПЛТ	обзорно-прицельная	+	-	-	-	-	-	-	+	+	2	МКИО	3.1.1	КД ГД
СОВИ	обзорно-прицельная	+	-	-	-	+	-	-	+	+	2	RS-485 ** RS-232 МКИО	3.1.2	О
БТИ	обзорная	+	-	-	+	+	-	-	-	-	-	RS-485	1.10	О ₁
БТИ-1	обзорно-прицельная	+	-	-	+	+	-	+	+	+	2	RS-485	1.10	О ₁
АТТ	обзорно-прицельная	+	-	+	+	+	-	-	+	-	1	RS-485 МКИО ARINC **	3.1.2	О ₁
АСТ	обзорно-прицельная	+	-	-	+	+	-	+	+	-	1	RS-485 МКИО	2.1.1 2.2.1	О ₁
ОВИ-МЗ(М2)	обзорно-прицельная	+	+	-	-	-	+	+	+	+	1-8	RS-485	1.7.1	О ₁ (КД ГК)
ТТА	обзорно-прицельная	+	-	-	-	+	+	+	+	+	2	-	1.7.1	О
СОВИ-ШТУРМ	обзорно-прицельная	+	-	-	+	+	-	-	+	-	2	RS-485	1.4.1	О
ВУВИ-АГ	поисковая	+	+	+	+	+	-	-	+	+	2	RS-232 *** RS-485	В5	КД ГК
БУВИ-01, -07, -08	поисковая	+	-	-	+	+	-	-	-	-	-	RS-485 RS-232	-	КД ГК
МКУВИ	вождения	+	+	+	+	+	+	+	+	+	1.5.4	ITU-R **** BT.654-4	1.5.4	О ₁
БОВИ	обзорно-прицельная	+	+	+	+	+	+	+	+	+	1.3	R-422	1.3	О ₁
МЦОВИ	поисковая	+	+	-	+	+	-	-	-	-	-	RS-232	-	КД ГК

* - полный аналоговый входной/выходной видеосигнал для черно-белого телевидения по ГОСТ 7845-92

** - Fibre channel

*** - полный аналоговый цветной входной/выходной видеосигнал по ГОСТ 7845-92

**** - для выходных видеоданных.

Обозначения в таблице: УВ –улучшение видения; К –комплексирование; ЭС –электронная стабилизация; КК – компенсация крена; ЭМ –электронное масштабирование; АО –автоматическое обнаружение; АЗ –автоматический захват; ЭСл –электронное слежение (в поле зрения); АСЗК –автоматическое сопровождение в замкнутом контуре; Ц –число сопровождаемых целей.

На настоящий момент времени произведено около двухсот пятидесяти изделий, они успешно эксплуатируются, выполняя функции цифровой обработки изображений в наземных, морских и воздушных объектах ВВТ. Продолжаются работы по их дальнейшему совершенствованию как в части алгоритмического, так и аппаратного обеспечения в целях развития функциональности и повышения эффективности решаемых задач.

Литература

1. Зеленюк Ю.И., Алпатов Б.А., Костяшкин Л.Н., Романов Ю.Н. Видеокомпьютерные технологии в разработке обзорных и обзорно-прицельных систем// Мир Авионики, №1-2, 2000, с. 28-32.
2. Травникова Н.П. Эффективность визуального поиска. - М.:Машиностроение, 1985.
3. Орлов В.А., Петров В.И. Приборы наблюдения ночью и при ограниченной видимости.- М.: Воениздат, 1989.
4. Богданов А.П., Романов В.Ю., Романов Ю.Н. Комплексирование телевизионных и тепловизионных видеоизображений. Материалы 11-ой международной научно-технической конференции «Проблемы передачи и обработки информации в сетях и системах телекоммуникаций», Рязань, 2002, с.41-43.
5. Зеленюк Ю., Костяшкин Л., Романов Ю., Давиденко В., Кузнецов С. Комплексирование разносектральной информации в приборе механика водителя// Военный парад, №6, 2006, с.40-42.
6. Алпатов Б.А., Селяев А.А. Алгоритм оценки местоположения объекта на двумерном изображении// Приборостроение, №5, 1988, с.21-27.
7. Алпатов Б.А., Бабаян П.В., Балашов О.Е., Степашкин А.И. Методы автоматического обнаружения и сопровождения объектов. Обработка изображений и управление. - М.: Радиотехника, 2008.
8. Алпатов Б. А., Блохин А.Н. Модели и алгоритмы обнаружения и выделения движущихся фрагментов изображений// Автометрия, №4, 1995, с.45-49.
9. Муравьев В.С., Муравьев С.И. Пространственно-временной алгоритм выделения объектов, наблюдаемых на облачных фонах // Цифровая обработка сигналов и ее применения. Тез. докл. 9-й междунар. конф. Том 2. – М.: 2007. – С. 375-377.

“OKHOTNIK” FAMILY OF MULTIPURPOSE VIDEOIMAGE PROCESSING SYSTEMS

Alpatov B.A., Blokhin A.N., Kostyashkin L.N., Romanov Y.N., Shapka S.V.

The paper presents the tasks and algorithmic support of “Okhotnik” videoimage processing system vision enhancement, automatic acquisition, target lock-on and tracking. It offers the vision enhancement operator algorithms, the algorithms of videoimage fusion implementation, the algorithms of different-type object selection against diversified background and the algorithm of line-of-sight control. The peculiarities of family system computing platforms in real time are also being analyzed in the paper. The developed modifications of “Okhotnik” family are also described here.

У в а ж а е м ы е а в т о р ы !

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

1) 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.*

НОВЫЕ ТЕХНОЛОГИИ ЦИФРОВОГО ТЕЛЕВИДЕНИЯ

Вилкова Н.Н., Генеральный директор ЗАО «МНИТИ», к.т.н.

Евстигнеев В.Г., начальник отдела ЗАО «МНИТИ», д.т.н., evstigneev_mniti@mail.ru, (495) 460-04-88

Зубарев Ю.Б., советник Генерального директора ЗАО «МНИТИ», д.т.н., член-корр. РАН

Минаев В.Н., директор Департамента радиоэлектронной промышленности Минпромторга России, д.т.н.

Ключевые слова: декорреляция, матрица, стандарты H.261, H.263, MPEG-1, MPEG-2, MPEG-4, H.264, JPEG-2000

Большинство стандартов видеокодирования, реализованных после 1990 года, основаны на некоторой единой модели видеокodeка, которая включает оценку движения и компенсацию высоких частот (обозначается DPCM – дифференциальная кодово-импульсная модуляция), стадии преобразования, квантования и энтропийного сжатия. Такая конструкция называется гибридной моделью DPCM/DCT [1]. К стандартам, работающим по указанной схеме, относятся H.261, H.263, MPEG-1, MPEG-2, MPEG-4, H.264.

Таким образом, основные кодеки видеоданных состоят, как правило, из трёх самостоятельных этапов (стадий): межкадровая декорреляция (декорреляция во времени), внутрикадровая декорреляция (декорреляция в пространстве) и энтропийное сжатие. Причём, основополагающие принципы реализации каждого из указанных трёх этапов компрессии-декомпрессии могут быть самые разные, важно чтобы они давали соответствующий положительный результат. Причем различные варианты реализации каждого из этапов взаимозаменяемы. Последнее обстоятельство согласуется с утверждениями Яна Ричардсона [1].

Проблему хорошего энтропийного сжатия к настоящему времени можно считать практически решённой. Так, адаптивное побитовое контекстно-ориентированное арифметическое кодирование в стандарте JPEG-2000 позволяет сжимать данные до предельной величины, которая рассчитывается исходя из энтропии этих данных.

Этого нельзя сказать об этапе пространственной декорреляции, о путях реализации которого по настоящее время продолжают споры. На сегодняшний день существует множество самых разных методов и подходов к решению этой задачи. Наиболее распространённые основаны на преобразовании исходных данных, заключающемся в переводе их из пространственной области в пространственно-частотную. Первый подход к решению этой проблемы был применён в стандартах JPEG, MPEG-2, MPEG-4, H.264. Пространственная и частотная составляющие этапа преобразования в указанных стандартах осуществляются раздельно друг от друга: пространственная реализуется путём разбиения исходного

Представлен аналитический обзор классических и современных подходов к решению задачи декорреляции цифровых изображений, проведен сравнительный анализ методов декорреляции цифровых изображений, применяемых в стандартах H.261, H.263, MPEG-1, MPEG-2, MPEG-4, H.264, JPEG-2000: констатировано, что возможности ДКП в решении задачи декорреляции цифровых изображений практически исчерпаны и, что решение задачи декорреляции цифровых изображений любых размеров необходимо искать в применении новых технологий цифрового телевидения, основанных на использовании декоррелирующих ортогональных целочисленных матриц.

изображения на квадратные подобласти (блоки), а частотная – с помощью дискретного косинусного преобразования (ДКП) значений пикселей внутри этих блоков. Очевидно, что данный подход явно искусственный.

Появившийся позднее кратно-масштабный анализ (КМА) и тесно связанные с ним вейвлет-преобразования предложили свой подход к решению задачи декорреляции видеоизображений. Обе составляющие пространственно-частотного преобразования реализуются одновременно в рамках единой методологии. Данный подход, в частности, использовали создатели стандарта JPEG-2000. Этот подход уже можно считать естественным и органичным.

Стандарт JPEG-2000 отдал предпочтение второму подходу. Однако создатели стандартов сжатия видеоданных, таких как MPEG-2, MPEG-4, H.264, пошли путём улучшения и модернизации первого подхода, основанного на ДКП.

Преобразование ДКП – это частный случай дискретного преобразования Фурье (ДПФ). ДПФ – универсальный метод, который впервые был применён для перевода данных (изображение, звук) из пространственно-временной области в частотную. Его недостатками являются избыточность и высокая вычислительная сложность. Однако такая сложность обеспечивает резервы для его модернизации и упрощения.

При реализации пространственной составляющей преобразования, основанного на ДКП, имеются трудности. В частности, существует ограничение, налагаемое на размеры блоков разбиения. Увеличение числа пикселей в блоке приводит к уменьшению корреляционной зависимости между ними и, следовательно, ухудшению результатов декорреляции при кодировании. Из литературы [2] известно, что на практике верхней границей величины блоков разбиения является размерность 8×8 .

Именно такие размеры блоков можно видеть в большинстве методов, основанных на ДКП. Считается, что декорреляция блоков, превосходящих по размерам данную величину, уже не имеет смысла. Но в этом случае данная размерность - наихудшая из возможных. Наилучшей будет та, где корреляция между пикселями максимальна, то есть 2×2 .

Создатели стандарта H.264 подвергли метод, основанный на ДКП, значительной модернизации, достигнув в этом процессе действительно больших успехов. Они уменьшили вдвое (по длине и высоте) размеры блоков, внутри которых производится ДКП, что повысило корреляцию между преобразуемыми пикселями. Помимо этого, им удалось реализовать ряд вычислений над числами в целочисленном формате, которые ранее представлялись в форме с плавающей запятой. Некоторые целочисленные элементы матрицы ДКП были заменены ими на приближённые. Полученный новый метод уже нельзя назвать методом ДКП, хотя он работает по принципу ДКП, будучи его хорошей модернизацией. Положительные результаты, полученные создателями стандарта H.264 в смысле минимизации вычислительной сложности этапа пространственной декорреляции, основанном на принципе ДКП, очевидны. Однако, как отмечают многие, данный стандарт отличается отсутствием гибкости. И, вероятно, дальнейшая оптимизация этапа пространственной декорреляции на основе ДКП уже невозможна. Требуется искать новые подходы к решению данной проблемы. Несмотря на это, как считает Ян Ричардсон [1], реклама стандарта H.264 полна оптимизма и сулит ему долговременный успех в будущем.

Создателям будущих стандартов при реализации этапа декорреляции в кодеках видеоданных вероятно придётся уже полностью отказаться от попытки решить эту проблему на основе ДКП. Замена искусственного метода перевода исходных данных в пространственно-частотную область на естественный уже происходит. Вместо морально устаревающих ДКП всё больше используются дискретные вейвлет-преобразования (ДВП), возникшие на основе теории кратно-масштабного анализа. Преобразования ДВП успешно применены в стандарте JPEG-2000 и дали хороший результат как по степени сжатия, так и по качеству восстановленного изображения.

Достоинства методов пространственной декорреляции изображений, построенных на основе теории кратно-масштабного анализа (пирамидальные преобразования, субполосное кодирование, вейвлет-преобразования и т.д.), в настоящий момент широко известны и описаны в литературе.

К их недостаткам следует отнести, прежде всего, отсутствие универсальности. Большинство этих методов требуют обязательного равенства длины и высоты исходного изображения натуральной степени числа два (дихотомический принцип). На практике при обработке видеоизображений, размеры которых отличаются от натуральных степеней числа два, пользуются искусственными приёмами, приводящими к повышению затратности процесса декорреляции. Это – существенное ограничение в выборе формата исходного изображения. В малочисленной литературе можно найти примеры крат-

но-масштабного трёхполосного кодирования [3].

Решение же задачи декорреляции исходных изображений естественным методом любых заданных размеров в рамках единой идеологии необходимо искать в применении декоррелирующих ортогональных матриц произвольного размера. Теория субполосного кодирования допускает это, но на практике такой подход почему-то не находит применения. Большинство кратно-масштабных методов декорреляции строятся по дихотомическому принципу, редко по трихотомическому. В основе многих дихотомических методов явно или неявно лежит классическая матрица Адамара, размером 2×2 . Декоррелирующие матрицы больших размеров строятся, как правило, методом ДКП или близкими к нему методами, оперирующими трансцендентными (тригонометрическими) функциями. Помимо высокой вычислительной сложности это приводит ещё к существенному уменьшению пространственной корреляции пикселей по мере увеличения размеров декоррелирующей матрицы. Последнего можно избежать, если декоррелирующими матрицами переводить значения пикселей внутри блока разбиения не в частотную (как ДКП), а в пространственно-частотную область (как ДВП). Матрица Хаара произвольного размера, равного натуральной степени числа два, например, даёт хорошую степень декорреляции, несмотря на то, что её размеры многократно превосходят «предельно допустимое значение» 8×8 . Описаний простых в применении декоррелирующих систем многополосного кодирования, охватывающих произвольные размеры исходного изображения, и схем генерации таких систем, применяемых на практике при работе с реальными видеоданными, в современной литературе не встречается.

Так обстоят дела на практике. Теория же говорит об обратном, давно существуют обобщения методов Адамара и Хаара, являющиеся типичными дихотомическими преобразованиями, позволяющие строить ортогональные (или унитарные, если речь идёт о комплексной области) матрицы и не только дихотомическими методами.

Матрицей Адамара порядка $N=2^n$ называется квадратная матрица размером $N \times N$ с элементами ± 1 , так, что

$$H_N \times H_N^T = N \cdot I_N, \quad (1)$$

где I_N – единичная матрица, а T – знак транспонирования [4].

Матрицу Адамара порядка N можно построить рекурсивно по схеме [4]:

$$H_N = \begin{bmatrix} H_{N/2} & H_{N/2} \\ H_{N/2} & -H_{N/2} \end{bmatrix}, \quad (2)$$

Приведённое выражение можно записать в терминах матричных операций [5]:

$$H_N = H_2 \otimes H_{N/2}, \quad (3)$$

где \otimes – символ так называемого кронеккеровского (или тензорного) произведения, которое определяется следующим образом:

Если A и B – квадратные матрицы соответственно порядков m и n , то кронеккеровское произведение матриц обозначается символом \otimes и записывается в виде

$A \otimes B$. Элементы матрицы, получившиеся в результате кронеккеровского умножения матриц A и B , равны $a_{ij}B$, то есть

$$A \otimes B = \{a_{ij}B\}. \quad (4)$$

Матрица H_2 в (3) – простейшая матрица Адамара, размером 2×2 , имеющая вид:

$$H_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}. \quad (5)$$

Как видно из определения кронеккеровского произведения, матрицы-сомножители не обязательно должны быть квадратными, а могут иметь произвольные прямоугольные размеры.

Из приведённой схемы построения матриц Адамара больших порядков видно, что их размеры обязательно должны выражаться натуральными степенями числа два. Существуют теоретические обобщения, которые позволяют отойти от этого требования. Это - так называемые матрицы Виленкина. Последняя строка матрицы H_2 в (5) является множеством двух корней квадратных из единицы. Если вместо этого взять три корня кубических из единицы, можно построить матрицу H_3 , размерами 3×3 , имеющую вид:

$$H_3 = \begin{bmatrix} 1 & 1 & 1 \\ 1 & j & j^2 \\ 1 & j^2 & j \end{bmatrix}, \quad (6)$$

где $j = \sqrt[3]{1}$ или по формуле Эйлера:

$$j = e^{2\pi/3i} = \cos 2\pi/3 + i \cdot \sin 2\pi/3, \quad (7)$$

Матрицы, указанного типа, именуется матрицами дискретных экспоненциальных функций (ДЭФ) [5], и, объединяясь друг с другом с помощью операции кронеккеровского произведения, образуют систему матриц Виленкина, каждая из которых является ортогональной (для области комплексных чисел – унитарной).

Классические матрицы Хаара строятся также по рекуррентной схеме [5]:

$$A_{2N} = \begin{bmatrix} A_N & \otimes & \begin{bmatrix} 1 & 1 \end{bmatrix} \\ I_N & \otimes & \begin{bmatrix} 1 & -1 \end{bmatrix} \end{bmatrix}, \quad (8)$$

исходная матрица A_2 – это уже упомянутая матрица Адамара, размером 2×2 , имеющая вид (5). A_N в (8) – текущая матрица Хаара, размером $N \times N$, где $N=2^n$ (n – номер текущей итерации), A_{2N} в (8) – новая матрица Хаара, размером $2N \times 2N$, полученная на текущей итерации, I_N в (8) – единичная матрица размером $N \times N$.

Для $N=2$, например, матрица I_2 имеет следующий вид:

$$I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \quad (9)$$

Из итерационной формулы (8) видно, что матрицы-строки, умножаемые справа кронеккеровским произведением на матрицы A_N и I_N , являются строками матрицы Адамара H_2 , которые следуют одна за другой в их естественном порядке.

Из указанной итерационной формулы (8) также видно, что размеры матриц Хаара могут равняться только

натуральной степени числа 2.

Итерационная формула (8) формирования матриц Хаара легко обобщается и на многомерный случай. В основе указанной итерационной формулы лежит матричная операция, известная в литературе [5], которую условно называют кратно-масштабным объединением (КМО). Эта операция определяется следующим образом.

Пусть даны две квадратные ортогональные (или унитарные) матрицы A – размером $p \times p$ и B – размером $q \times q$. Кратно-масштабным объединением этих матриц будем называть матрицу, полученную следующим образом:

$$A \circ B = \begin{bmatrix} A \otimes (B)_0 \\ I_p \otimes (B)_1 \\ I_p \otimes (B)_2 \\ - & - & - \\ I_p \otimes (B)_k \\ - & - & - \\ I_p \otimes (B)_{q-1} \end{bmatrix}, \quad (10)$$

где A – первая матрица кратно-масштабного объединения, $(B)_k$ – k -я строка второй матрицы кратно-масштабного объединения (то есть матрицы B), $k = 0, 1, 2, \dots, q-1$,

I_p – единичная матрица, размером $p \times p$,

\otimes – символ кронеккеровского произведения,

\circ – условный символ кратно-масштабного объединения двух ортогональных (или унитарных) матриц.

Размер матрицы-результата (10) будет равен произведению размеров объединяемых матриц (A и B), то есть $(pq) \times (pq)$.

Если матрицы A и B относятся к системе матриц ДЭФ, то их кратно-масштабное объединение будет одной из комплексных матриц Хаара (другими словами матриц Виленкина-Хаара).

Схема формирования матриц Виленкина (или матриц Виленкина-Хаара) состоит из многократно повторяющегося действия (кронеккеровского произведения или кратно-масштабного объединения) над множеством элементарных матриц (в данном случае - матриц ДЭФ).

Матрицы ДЭФ строятся из комплексных чисел. Работать с ними неудобно. Если заменить их на целочисленные ортогональные матрицы [6], размеры которых являются простыми числами натурального ряда, и известен алгоритм генерации всех целочисленных ортогональных матриц с указанными размерами, то пользуясь схемой построения матриц Виленкина и Виленкина-Хаара, заменив лишь базисные матрицы с комплексных на целочисленные, можно строить целочисленные ортогональные матрицы произвольных размеров.

Классические матрицы Адамара, размерами равными произвольной натуральной степени числа два, строятся по итерационной схеме (3), путём кронеккеровского (тензорного) умножения матрицы, полученной на предыдущей итерации, на элементарную матрицу Адамара, размером 2×2 (5), которая одновременно является исходной в итерационном процессе.

Классические матрицы Хаара строятся аналогично, только вместо кронеккеровского произведения используется операция КМО.

Обе эти операции обладают следующими свойствами:

1) объединение с их помощью двух ортогональных декоррелирующих матриц даёт в результате новую ортогональную декоррелирующую матрицу; если первые строки объединяемых матриц состоят из одних единиц, то и первая строка составной матрицы состоит из одних единиц;

2) обе операции некоммутативны, что открывает возможность генерации различных ортогональных декоррелирующих матриц одного и того же размера. Такая многовариантность создаёт возможность оптимального выбора наиболее подходящей для решения данной конкретной задачи матрицы.

3) размер ортогональной декоррелирующей матрицы-результата равен произведению размеров двух объединяемых более малых ортогональных декоррелирующих матриц.

Если размер исходного изображения по длине выражается целым составным числом, то, согласно «основной теореме арифметики», его можно представить в виде произведения простых меньших по величине чисел в целых неотрицательных степенях. Следовательно, согласно третьему свойству, декоррелирующая матрица, необходимая для одномерного преобразования исходного изображения по длине, может быть построена из набора элементарных матриц, размеры которых выражаются соответствующими простыми числами. Матрицы для вертикального преобразования строятся по аналогичной схеме. Для двумерного преобразования необходимы две матрицы, одна для преобразования по горизонтали, другая - по вертикали.

Второе свойство хорошо использовать для передачи конфиденциальной информации. Для правильного декодирования сжатого изображения необходимо знать порядок и структуру элементарных матриц, из которых построена декоррелирующая матрица. Это является как бы ключом, не зная которого затруднительно получить доступ к закодированной видеoinформации.

В середине прошлого века советскими учёными Н.Я. Виленкиным, Л.С. Понтягиным и др. была разработана чёткая система генерации ортогональных (унитарных) матриц произвольных размеров, элементы которых выражаются комплексными числами. Использование таких систем на практике – затруднительно. Для устранения этого недостатка необходима замена способа формирования элементарных матриц. Требования к этим матрицам должны быть следующими: они должны быть ортогональными и декоррелирующими; их элементы должны быть целочисленными; их размеры должны равняться простым числам (причём метод их генерации должен охватывать всю совокупность простых чисел их размеров); они должны переводить исходные данные в пространственно-частотную область, то есть быть «хааровскими», а не «адамаровскими» (следовательно, построение такой системы на основе ДКП и близких к нему методов неприемлемо); первая строка всех этих матриц должна состоять из одних единиц.

Методы формирования системы ортогональных матриц с заданными свойствами, разработанные в ЗАО «МНИТИ» [5], являясь первым этапом новых технологий цифрового телевидения, позволяют отказаться от традиционного ДКП – декорреляции видеоизображений и перейти к технологии матричной декорреляции. Новые технологии декорреляции цифровых видеоизображений просты и универсальны и могут заменить существующие методы реализации этапа пространственной декорреляции в кодеках всех стандартов цифрового телевидения. Положительные результаты такой замены очевидны: работа с целыми числами, обработка изображений произвольных размеров и форматов в рамках единой идеологии без искусственных приёмов, многовариантность как основа выбора оптимального способа преобразования изображений в соответствии с конкретно поставленной задачей.

В настоящее время в ЗАО «МНИТИ» ведётся работа по патентованию новых алгоритмов и способов декорреляции цифровых изображений для дальнейшего их использования при модернизации кодеков в стандартах MPEG-2, MPEG-4 и H.264. Это может обеспечить приоритет запатентованной российской системе при работе с цифровыми телевизионными изображениями произвольных форматов с использованием отечественной электронной компонентной базы, а также уменьшить требования к производительности БИС, реализующих кодеки стандартов MPEG-1, MPEG-2, MPEG-4, H.261, H.263 и H.264 на 24÷50 %.

Более детальное изложение результатов использования новых технологий в решении задачи декорреляции цифровых телевизионных изображений можно найти в публикациях [6, 7, 8].

Выводы

В исследованиях проводимых в ЗАО «МНИТИ» по проблеме новых технологий цифрового телевидения решаются следующие задачи:

- оценка сложности и затратности известных алгоритмов декорреляции цифровых телевизионных изображений;
 - разработка и анализ новых алгоритмов формирования целочисленных ортогональных декоррелирующих матриц;
 - разработка и анализ новых алгоритмов декорреляции цифровых телевизионных изображений на основе целочисленных ортогональных декоррелирующих матриц;
 - оценка сложности новых алгоритмов декорреляции цифровых телевизионных изображений;
 - программное моделирование новых методов и алгоритмов декорреляции цифровых телевизионных изображений;
 - патентование матричных способов декорреляции цифровых изображений и способов формирования целочисленных ортогональных матриц различного назначения, размеров и декоррелирующих качеств;
 - обеспечение конфиденциальности цифровых телевизионных изображений некриптографическими методами, передаваемых по незащищенным каналам связи.
- Сопоставление отечественных матричных способов декорреляции цифровых изображений со способами

декорреляции, применяемыми в стандарте H.264, защищенными патентами США, показывает преимущества разработок ЗАО «МНИТИ», требующих производительность процессорной БИС, по крайней мере, от 24 до 50 % меньше, чем в стандарте H.264.

Литература

1. Ян Ричардсон «Видеокодирование. H.264 и MPEG-4 – стандарты нового поколения», ТЕХНОСФЕРА, Москва, 2005 год
2. Д. Сэломон. Сжатие данных, изображений и звука. – ТЕХНОСФЕРА, Москва, 2004.
3. А.В. Гельманшин, В.П. Дворкович «Новый подход к использованию вейвлет-фильтров при обработке изображений», Цифровая обработка сигналов № 1/2008, Москва, стр. 37-42.
4. Э.Е. Дагман, Г.А. Кухарев. Быстрые дискретные ортогональные преобразования. – Издательство «НАУКА», Сибирское отделение РАН Новосибирск, 1983.
5. В.Г. Евстигнеев, А.В. Бондаренко, В.Д. Лебедев. «Формирование ортогональной системы преобразования изображений с размерностью, отличной от натуральной степени числа два», журнал «Вопросы оборонной техники». Научно-технический сборник, серия 3, выпуск 6 (394), 2008 г., Москва, стр. 39-50.
6. В.Г. Евстигнеев, А.В. Бондаренко, А.Н. Кошарновский, В.Д. Лебедев «Современные подходы к решению задачи декорреляции цифровых изображений» «Техника средств связи», серия «Техника телевидения», юбилейный выпуск, 2010 г., Москва, стр. 26.
7. В.Г. Евстигнеев, А.В. Бондаренко, А.Н. Кошарновский, В.Д. Лебедев «Дихотомические и недихотомические методы декорреляции цифровых изображений». «Системы и средства связи, телевидения и радиовещания». ISSN 2079-6137 Выпуск 1,2 2010 г., Москва., стр. 44.
8. В.Г. Евстигнеев, А.В. Бондаренко, А.Н. Кошарновский, В.Д. Лебедев «Оценка вычислительной сложности решения задачи декорреляции цифровых монохромных изображений». «Системы и средства связи, телевидения и радиовещания». ISSN 2079-6137 Выпуск 1,2 2010 г., Москва, стр. 50.



TEXAS INSTRUMENTS INC. выпускает новый процессор платформы **DaVinci TMS320DM8168**. Процессор открывает новые горизонты вычислительных мощностей, обеспечивая 4-кратное увеличение производительности в формате HD, реализацию перспективных кодеков и видеоакселераторов, обладает высокой степенью интеграции при низком энергопотреблении.

Процессор строится на базе ядра общего назначения ARM Cortex A-8 с частотой 1 ГГц и DSP-ядра с плавающей точкой C674x+ с частотой 1 ГГц.

Подробнее: <http://www.ti.com>

TEXAS INSTRUMENTS INC. представляет новую линейку микроконтроллеров **TMS570** для систем безопасности транспортных средств. Контроллеры используют процессорное ядро ARM® Cortex™-R4F с плавающей точкой, работающее на частоте 160 МГц.

Подробнее: <http://www.ti.com/ww/en/mcu/tms570/index.shtml>



TEXAS INSTRUMENTS INC. предоставляет разработчику систем на базе DSP возможность использования нового бесплатного программного инструментального средства **C6EZFlo** графической разработки приложений для процессоров C6000™.

Подробнее: <http://www.ti.com/c6flo-etechtf>

УДК 681.3.053

ПЕРСПЕКТИВЫ РАЗВИТИЯ ВИДЕОКОНФЕРЕНЦСВЯЗИ В РОССИИ

*Дворкович В.П., зам. директора ФГУП «ГРЧЦ» д.т.н., dvp@nircom.ru
Фёдоров В.Ф., зам. начальника отдела ФГУП «ГРЧЦ», к.м.н.*

Ключевые слова: видеоконференцсвязь, телекоммуникации, технические средства, программные решения.

Несмотря на появление первых видеотелефонов ещё в 1964 г. (Videophone от Bell Laboratories), началом массового распространения видеоконференцсвязи, наверное, следует считать вывод в декабре 1996 г. на рынок совместного продукта фирм 3Com и PictureTel [1]. Начав с относительно простых систем, PictureTel сравнительно быстро вышла на производство терминалов групповой ВКС, поддерживающих многоточечные видеоконференции. Вслед за оборудованием PictureTel появилась разработка ProShare Intel, которая не претендовала на изысканность, однако обладала достаточной функциональностью для организации конференций «точка - точка». Некоторое время спустя фирма Intel отказалась от этого сектора рынка. В течение следующих пяти лет на мировой и российский рынки свою продукцию в области видеоконференцсвязи вывели фирмы Polysom, Sony, VCON, VTEL, Aethra.[2]. Несколько позже других на рынке появилась норвежская фирма Tandberg, однако на европейском и российском рынках она за короткое время вышла в лидеры как по объёмам поставок, так и по ассортименту предлагаемых продуктов.[3]

В России технологии ВКС появились несколько позже, чем в странах Америки и Европы, однако отставание не было фатальным. Так уже в 1997 году началась реализация телемедицинского проекта «Москва-регионы России», в основе которого – дистанционное консультирование с применением ВКС ведущими столичными специалистами своих коллег из регионов по сложным клиническим случаям [4]. К настоящему времени проектом, которым руководит Российская ассоциация телемедицины, охвачено более 130 телемедицинских центров и пунктов в большинстве российских регионов. В 1999 году в судебной системе Российской Федерации впервые в мировой практике началось внедрение технологий ВКС для проведения судебных процессов. За десять лет в судах общей юрисдикции и СИЗО было установлено более тысячи терминалов ВКС, с помощью которых уже рассмотрено более 350 тысяч уголовных дел [5].

Последние пять-семь лет на российском рынке телекоммуникационных услуг можно охарактеризовать как период бурного роста практического интереса к системам видеоконференцсвязи. Достаточно посмотреть публикации по этой теме в Интернете, чтобы убедиться, что эта технология год за годом всё шире внедрялась в бизнес и во всё новые сферы общественной жизни [6, 7].

Однако большинство публикаций или напрямую оплачены производителями и продавцами конкретного оборудования ВКС, что хорошо видно из предпочтений автора при описании возможных решений, или написаны с позиций

Видеоконференцсвязь (ВКС), по сравнению с другими средствами телекоммуникаций, является относительно молодой технологией, но в тоже время её использование становится рутинным для всё более широкого круга пользователей. В статье подробно излагаются проблемы реализации видеоконференцсвязи в России.

анализа рынка.

В то же время точки зрения **потребителя** средств ВКС и **отечественных разработчиков** в этой сфере остаются вне сферы внимания профильной прессы и сетевых информационных ресурсов. Авторы настоящей публикации попытались восполнить этот пробел.

Пользователи реальные и потенциальные

Любая новая техника и технология предназначены, по мнению разработчиков и производителей, для использования определённым кругом лиц. Как же позиционируется ВКС сегодняшним рынком?

Прежде всего – это бизнес. Крупные компании, имеющие распределённую структуру и филиалы на разных континентах, безусловно, выигрывают в оперативности управления и экономят средства при широком внедрении ВКС за счёт сокращения командировочных расходов и своевременности принятия бизнес-решений [8].

Далее – госструктуры. Во-первых, силовые ведомства, для которых оперативность управления едва ли не синоним успешности выполнения операций [9, 10].

Во-вторых, службы, ликвидирующие последствия чрезвычайных ситуаций [11, 14]. Как пишет автор названного источника: «Использование мобильных ВКС помогает МЧС быстро развертывать специализированную инфраструктуру при масштабных авариях или природных катастрофах, что обеспечивает взаимодействие различных служб, которые работают в подобных случаях».

В-третьих, судебная и пенитенциарная системы [5]. Применение ВКС при повторном рассмотрении дел заключённых или при рассмотрении новых дел позволяет отказаться от этапирования осуждённого из удалённых мест лишения свободы к месту проведения судебного заседания, а также привлечь к участию в нём свидетелей, находящихся в других населённых пунктах вне зависимости от их удалённости [12].

В-четвёртых, в последнее время значительное количество совещаний центр-регионы проводятся высшим руководством страны в режиме ВКС. Губернаторы и другие служащие докладывают президенту или премьеру о состоянии дел в регионах и ходе выполнения поручений руководства [13, 14].



В-пятых, руководство регионального уровня также переходит к проведению территориально распределённых совещаний, вместо вызова подчинённых непосредственно в региональный центр [15].

Всё более значимую роль играет видеоконференцсвязь в медицине [16, 17, 18]. Телемедицинские консультации, консилиумы, тематические дистанционные циклы повышения квалификации, отдельные лекции проводятся всё большим числом ведущих клинических учреждений столицы и региональных центров. На сегодня только в Москве с различной интенсивностью работают более 20 консультирующих телемедицинских центров и пунктов.

Дистанционное образование с применением видеоконференцсвязи внедряется многими вузами страны. В вузах, имеющих филиалы в других городах (регионах), как правило, вначале ВКС используется для совещаний руководящего состава и преподавателей, повышения квалификации и тренингов сотрудников [19], а затем – для чтения лекций, проведения семинаров и даже для сдачи зачётов и экзаменов [20].

Из многочисленных публикаций в Интернете может сложиться впечатление об успешности внедрения технологий ВКС в мире вообще и в России в частности. Так ли это? Для ответа на этот вопрос достаточно ответить на два других:

1) Есть ли у большинства пользователей телефонной связи желание видеть собеседника при разговоре?

2) Каково количественное соотношение мобильных телефонов и терминалов видеоконференцсвязи?

На первый вопрос можно ответить коротко – да!

На второй ответить сложнее, т.к. реальных цифр по продажам терминалов ВКС найти не так просто. В то же время, очевидно, что если мы скажем, что это соотношение исчисляется порядками, то мы не ошибёмся. Даже если отнести к пользователям ВКС всех, кто пользуется «Скайпом» и аналогичными программами, всё равно речь будет идти о миллионах. Пользователей же мобильной связи – миллиарды. Если же учитывать только пользователей стандартных аппаратно-программных и программных терминалов ВКС, то соотношение будет не менее чем с пятью нулями.

Отдельно следует сказать о двух направлениях потенциально широкого внедрения технологий видеоконференцсвязи.

Во-первых, это – дистанционное образование. Не секрет, что в нашей стране вузы с сильным преподавательским составом распределены крайне неравномерно. А желание юных граждан из любого населённого пункта получить образование высокого качества вполне законно и должно только приветствоваться. Но не могут же престижные вузы вместить всех желающих, да и не все студенты могут платить за проживание в крупных городах. Наиболее простой выход видится только в распространении дистанционного обучения с применением технологий ВКС.

Во-вторых, это – домашняя телемедицина. Известно, что в развитых странах продолжается старение наций. С каждым годом относительное количество престарелых граждан с множеством хронических заболеваний растёт, а относительное количество медицинских работников остаётся тем же или уменьшается. При сохранении такой тенденции в скором времени поликлиническое звено здравоохранения будет просто парализовано. И снова – выход во внедрении технологий ВКС. Как показывает зарубежный опыт, регуля-

ное общение по видеосвязи медицинского работника среднего звена с пожилыми пациентами (телепатронаж) существенно увеличивает их приверженность к выполнению назначений и эффективность лечения. При этом медработник тратит на каждого пациента в среднем не более пяти минут в день, а нагрузка на поликлинику сокращается в разы (по разным данным от 2-х до 5-и раз).

Итак, вроде бы у рынка видеоконференцсвязи блестящие перспективы. Но почему же потенциальные пользователи ВКС не становятся реальными?

Структура технических средств ВКС

Минимальное оборудование, необходимое для использования видеоконференцсвязи, это:

- **терминал**, работающий хотя бы по одной из рекомендаций Международного союза электросвязи (МСЭ – англ. ITU-T) – H.320 или H.323;
- **модем**, работающий по той же рекомендации (или по обеим);
- подключенная **«последняя миля» сети** Интернет достаточной пропускной способности или канал (каналы) цифровой телефонии (сети ISDN).

Для того чтобы в видеоконференции могли участвовать три человека и более уже необходимо использовать **сервер многоточечной ВКС** (англ. Multipoint Control Unit - MCU), работающий в рамках тех же рекомендаций H.320 или H.323.

Многие производители технических средств ВКС заявляют о поддержке их изделиями перспективного протокола SIP, идущего на смену H.323, однако авторам не удалось найти описания сетей ВКС, реально работающих в России по этому протоколу.

Терминалы ВКС, представленные на современном рынке, это, как правило, моноблочный или модульный комплект из видеокамеры, дисплея и кодирующе-декодирующего устройства (кодека), выполняющие много полезных функций:

- Выбор удалённого абонента (адреса в IP-сети или номера в сети ISDN).
- Вызов удалённого абонента.
- Ответ на вызов удалённого абонента.
- Передача и прием звука (аудиоинформации).
- Передача и прием изображения (видеоинформации).
- Настройка параметров связи (IP-адрес, маски подсети и шлюза, полоса, разрешение и др.).
- Работа с адресной книгой (внесение, редактирование и удаление адресов и идентификаторов абонентов).
- Выбор источника (источников) локального видео (если терминал позволяет использовать больше одной камеры).
- Включения/отключения «картинки в картинке» (PiP).
- Выбор положения «картинки в картинке» (PiP) на экране (в одном из углов).
- Включение/отключение локального микрофона.
- Регулирование громкости входящего звука (динамиков).
- Регулирование громкости исходящего звука (чувствительности микрофона).
- Визуальная и звуковая индикация событий и ключевых состояний (канала связи, режима записи, отключения микрофона и т.п.).
- Переход в режим двух видеопотоков и обратно (если такая функция предусмотрена).

- Переход в режим защищённого сеанса связи и обратно.
- Переход в «спящий» режим при неактивности более установленного времени (например, 5 минут), активация при поступлении входящего вызова.
- Контроль параметров установленного соединения (полоса, скорость, частота кадров, протоколы, потери, джиттер и т.п.) и их отображение по запросу оператора.
- Возможность включения/отключения автоматического ответа на входящие вызовы.
- Индикация: входящих, исходящих, пропущенных (успешных, неуспешных) вызовов.

Для реализации описанного многообразия в кодеках встроены специализированные процессоры и «защиты» фирменные программные средства.

Большинство терминалов ВКС оснащены специальными PTZ – камерами (PTZ (англ.) – pan-tilt-zoom), позволяющими оператору (и/или удалённому абоненту) выбирать и увеличивать объект для отображения из окружающего пространства.

В ряде терминалов реализована функция автоматической наводки камеры на говорящего по голосу.

Многие терминалы имеют возможность передавать и/или принимать два видеопотока, причём второй из них – в форматах персональных компьютеров (в соответствии с рекомендациями МСЭ Н.239), однако эта функция, как правило, предоставляется опционально за дополнительную плату (~\$ 2000).

Некоторые терминалы имеют возможность работать из локальной сети сквозь NAT и FireWall в соответствии с рекомендациями МСЭ Н.460.18, Н.460.19.

К названным базовым функциям производители добавляют некоторое количество сервисных возможностей, повышающих привлекательность дорогостоящих изделий.

Модемы, используемые в ВКС, это устройства осуществляющие функцию оконечного оборудования линии связи, т.е. преобразователя сигналов между терминалом ВКС и канальным оборудованием провайдера.

Наиболее часто в России используются модемы двух типов:

ISDN — модемы для цифровых коммутируемых телефонных линий;

DSL — применяются для организации выделенных (некоммутируемых) линий, используя обычную телефонную сеть.

Существенно реже используются беспроводные модемы:

Спутниковые – модемы для сетей INMARSAT и VSAT.

3G – модемы сотовых сетей (например, SkyLink).

4G – модемы беспроводных сетей WiMax (например YOTA).

Ещё реже используются модемы телевизионных (коаксиальных) кабельных сетей в стандарте **DOCSIS**.

Передача аудио/видеопотоков между абонентами требует использования **каналов** гораздо большей пропускной способности, чем обычные телефонные.

На начальных этапах внедрения ВКС такими каналами служили линии цифровой телефонии ISDN, объединяемые в параллельный канал требуемой пропускной способности, в соответствии с рекомендациями МСЭ Н.320. Так для обмена потоками с разрешением CIF (Common Intermediate Format — стандарт размера изображения графических и

видеофайлов равный 352 x 288 пикселей) при частоте до 25 кадров в секунду обычно использовались 3 канала BRI (англ. BRI – *Basic Rate Interface* – обеспечивает пользователю предоставление двух цифровых В-каналов по 64 кбит/с и однополосный канал сигнализации D со скоростью передачи данных 16 кбит/с.). При этом пользователь фактически монополюбно занимал 6 телефонных каналов на время сеанса связи и оплачивал их по действующим тарифам телефонных компаний.

По мере повышения доступности широкополосного доступа к публичному Интернету всё чаще применяется видеоконференцсвязь по IP-каналам в соответствии с рекомендациями МСЭ Н.323. Пока оба варианта сосуществуют, но, учитывая тенденции развития IP-сетей и неразвитость сетей ISDN в России, можно предположить переход в ближайшем будущем к абсолютному доминированию IP варианта. На сегодня такое доминирование Н.323 в России сдерживается только низкой стабильностью отечественных IP-сетей и нежеланием многих провайдеров улучшать инфраструктуру и поддерживать гарантированное качество (QoS англ. *Quality of Service* — качество обслуживания. Этим термином в области компьютерных сетей называют вероятность того, что сеть связи соответствует заданному соглашению о трафике) [22].

Некоторые производители оборудования для ВКС в последнее время заявляют, что дни «обычной» видеоконференцсвязи сочтены, и в ближайшее время они планируют поставлять только оборудование для HD видеосвязи [23].

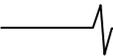
Если при сеансах обычной ВКС (352x288 пикселей при 25 кадрах в секунду) зачастую не хватает пропускной способности каналов связи (с учётом реальной загрузки множества маршрутизаторов по пути прохождения сигнала), то как должен транспортироваться видеопоток с разрешением 1940 x 1080 при тех же 25 кадрах в секунду?

Сервер, как правило, представляет собой сложный и дорогой аппаратно-программный комплекс, осуществляющий следующие функции:

- Вызов более одного абонента.
- Приём вызова более чем от одного абонента.
- Декодирование входящих аудио/видеопотоков от каждого из удалённых абонентов.
- Коммутация входящих видеопотоков (превращение их в единый поток с разделением экрана в соответствии с заданной раскладкой; в простейшем случае площадь экрана делится пропорционально количеству входящих потоков).
- Коммутация входящих аудиопотоков.
- Кодирование исходящего скоммутированного видеопотока.
- Кодирование исходящего скоммутированного аудиопотока.
- Отправка каждому из удалённых абонентов скоммутированного видеопотока.
- Отправка каждому из удалённых абонентов скоммутированного аудиопотока.

Кроме этого, серверы некоторых производителей могут варьировать раскладку изображений на одном или двух экранах, подключать к видеоконференции аудиоабонентов и т.п.

Существует альтернативная идеология построения серверов ВКС, когда сервер выступает в роли интеллектуального маршрутизатора. При этом резко снижаются требования к



производительности процессора сервера, т.к. ко всем участникам конференции просто перенаправляются два потока от тех абонентов, которые ведут диалог в настоящий момент. Естественно, что в этом случае в автоматическом либо ручном варианте определяется очерёдность участия в дискуссии, а требования к пропускной способности входящего к серверу канала возрастают пропорционально числу участников конференции.

Аппаратные и программные решения

Всё вышесказанное относительно терминалов ВКС касалось технических решений на базе специализированных процессоров (аппаратных кодеков). В то же время, на современном телекоммуникационном рынке достаточно предложений по «программным» терминалам ВКС. Кавычки применены, потому что реально, конечно, эти терминалы работают на базе персональных ЭВМ, т.е. используют аппаратные средства компьютеров и их операционную среду. Среди таких решений можно назвать продукцию фирм Polycom [24], Tandberg [25], VCON [26], Microsoft [27].

Если исходно программные решения в области ВКС существенно уступали аппаратным (например, программа NetMeeting от Microsoft обладала очень скромными возможностями по сравнению с аппаратными терминалами Polycom FX), то сегодняшние решения во всё возрастающей степени стирают грань различий в их возможностях с точки зрения пользователя. Так HD4000 от VCON уже позволяет передавать видео высокой чёткости и использовать протокол H.239 для передачи двух видеопотоков, что совсем недавно было прерогативой дорогостоящих групповых аппаратных систем ВКС высокого уровня. До настоящего времени, пожалуй, единственным серьёзным преимуществом аппаратных решений является проработка акустической части систем:

- высокоэффективное эхоподавление,
- высокоэффективное шумоподавление,
- автоматическое наведение камеры на говорящего.

В остальном программные системы с точки зрения пользователя становятся всё привлекательнее:

- они дешевле,
- они гибче (позволяют большую комбинаторику функций и аппаратных узлов),
- они динамичнее (обновления программных решений происходят чаще и не требуют замены аппаратных составляющих),
- они мобильнее (производительный ноутбук, веб-камеру высокого разрешения и аудиогарнитуру всегда можно взять с собой).

Альтернативные сервисы

Пользователь (извините за тавтологию!) хочет пользоваться определённым набором функций ВКС и совсем не хочет много платить за это. Идя навстречу его пожеланиям, был разработан и внедрён в сентябре 2003 года сетевой комплект сервисов Skype [28]. Комплект включает возможности:

- зашифрованной голосовой связи между компьютерами, подключенными к сети Интернет,
- зашифрованной аудиоконференции,
- пересылки файлов,
- обмена текстовыми сообщениями,
- зашифрованной видеоконференцсвязи «точка-точка».

Программа терминала предоставляется бесплатно, а

часть сервисов, являющихся платными, стоят дёшево.

Однако, как известно, «бесплатный сыр бывает только в мышеловке», да и то, только для второй мышки...

Во-первых, Skype не поддерживает протокола H.323, а, следовательно, несовместим с серверами и сетями стандартных систем ВКС.

Во-вторых, цитируем: «В отличие от многих других программ IP-телефонии, для передачи данных Skype использует P2P-архитектуру. Каталог пользователей Skype распределён по компьютерам пользователей сети Skype, что позволяет сети легко масштабироваться до очень больших размеров (в данный момент более 100 миллионов пользователей, 15-20 миллионов онлайн) без дорогой инфраструктуры централизованных серверов. Кроме того, Skype может маршрутизировать звонки через компьютеры других пользователей. Это позволяет соединяться друг с другом пользователям, находящимися за NAT или брандмауэром, однако создаёт дополнительную нагрузку на компьютеры и каналы пользователей, подключённых к Интернету напрямую» [28].

В-третьих, несмотря на заверения разработчиков о конфиденциальности связи, существуют данные о возможности перехвата сообщений [28], а использование протокола шифрования AES-256, который не сертифицирован на территории Российской Федерации, ставит российских пользователей вне закона.

В-четвёртых, Skype является коммуникационной средой для переноса вирусов, т.к. имеет встроенные эффективные алгоритмы прохождения сквозь NAT и брандмауэры [28].

В-пятых, Skype используется для кражи конфиденциальной информации с компьютеров пользователей, включая пароли для различных сервисов (например, входа в систему «электронный банк») [28].

На российском рынке активно продвигается система «Видеопорт» [29]. Однако и ей присущи неудобства других проприетарных систем: невозможность связи со стандартными терминалами ВКС, привязка к собственному серверу для расширения функциональности, бесплатность только минимального набора функций. Увеличить же функциональность можно только начав платить абонентскую плату (вместо разового платежа за лицензию, как это сделано у Sony или Tandberg).

Есть ещё ряд бесплатных программ ВКС, не поддерживающих стандарты МСЭ. Некоторые из них упрощены до предела (с точки зрения пользователя), например программа Logotech Vid, поставляемая вместе с веб-камерами высокого разрешения [29]. Однако, их функциональность, как правило, весьма ограничена, а отсутствие возможности установления связи со стандартным оборудованием ВКС делает их неудобными во многих случаях применения.

Цена решений

Когда речь заходит о широте охвата рынка, то многое определяется двумя факторами. На первом месте, конечно, цена изделия. Как бы потребителю не хотелось иметь некую вещь, но если в кошельке нет требуемой суммы и заработать её в ближайшем будущем не представляется возможным, он её не купит. Аналогично и с аппаратными терминалами ВКС. Выложить от шести до двадцати тысяч долларов (т.е. цену тех самых массовых автомобилей!) за возможность видеть, а не только слышать собеседника, готов далеко не каждый потребитель услуг связи.

Программные терминалы гораздо доступнее по ценам. Так, ещё два года назад, на российском рынке программный терминал «vPoint HD» фирмы Emblaze-VCON предлагался по ценам от 5200 р. до 18000 р., в зависимости от подключаемого набора функций, т.е. **ценовая разница с аппаратными решениями в десятки раз в пользу программных**.

Итак, какие же решения в области видеоконференцсвязи нужны пользователю?

Вначале о терминалах.

подавляющему большинству потенциальных пользователей нужны дешёвые решения при ограниченной функциональности. Часть потенциальных пользователей готова доплачивать за расширение функциональности, но их уже гораздо меньше. И лишь очень малое количество потенциальных пользователей готово много платить за имиджевые аппаратные решения, да и то, как правило, не своими, а государственными или корпоративными деньгами.

Теперь о серверах.

Покупателей дорогих аппаратных серверов многоточечной связи на рынке ничтожно мало. Большинство ограничивается серверами многоточки, встроенными в терминал. Ограничение в количестве от 4-х до 9-и одновременно подключаемых удалённых абонентов (что обеспечивают встроенные MCU), как правило, соответствуют реальным запросам потребителя. Однако при широком внедрении дистанционного образования и домашней телемедицины вопрос о серверах станет узким местом проектов с малым бюджетом. И именно тогда на первый план должны выйти программные серверы. Безусловно, программный сервер на большое число подключений, построенный по коммутирующему принципу, требует аппаратной платформы в виде компьютера очень высокой производительности. В то же время, его цена всё равно будет ниже, чем у аппаратных решений. Но ведь для большинства применений достаточно маршрутизирующего программного сервера многоточечной ВКС! Тем

более что пропускная способность «последней мили» большинства пользователей постоянно растёт из-за высокой конкуренции между провайдерами услуг широкополосного доступа в Интернет [31].

Могут ли что-то предложить своим пользователям отечественные разработчики, с учётом всего вышесказанного, и если «да», то что же они могут предложить?

Программные решения VPhone

Разработка алгоритмов и программ для передачи цифрового видеопотока по относительно узкополосным и нестабильным каналам связи была начата коллективом сотрудников НИИ Радио ещё в 1996 году. И уже через год разработка была представлена в Женеве вниманию международного сообщества. Программный комплекс под названием «VPhone» оказался интересен для проекта «уличного телевидения» во время второй избирательной кампании на Украине в 1999 году. Программные терминалы VPhone, установленные на десятках микроавтобусов, использовались для установления видеомостов между избирательным штабом и «улицей» во многих населённых пунктах Украины. Эффект «живого диалога с народом» внёс свой вклад в повышение рейтинга Президента и его победу на выборах [32].

Несколько позже программный комплекс VPhone привлёк внимание вузов, работающих в технологиях дистанционного образования: Современной гуманитарной академии [33] и Российского нового университета [35]. Наличие программного сервера многоточечной видеоконференцсвязи, одновременная передача видео/аудиоинформации и текстовых сообщений, возможность передачи двух и более видеопотоков, удобный инструментарий оператора видеосервера – всё это позволяет гибко конфигурировать многоточечную видеоконференцию в учебном процессе, охватывая в одном сеансе десятки филиалов и транслируя их слушателям лектора, презентацию и демонстрации опытов одновременно (рис. 1-2).

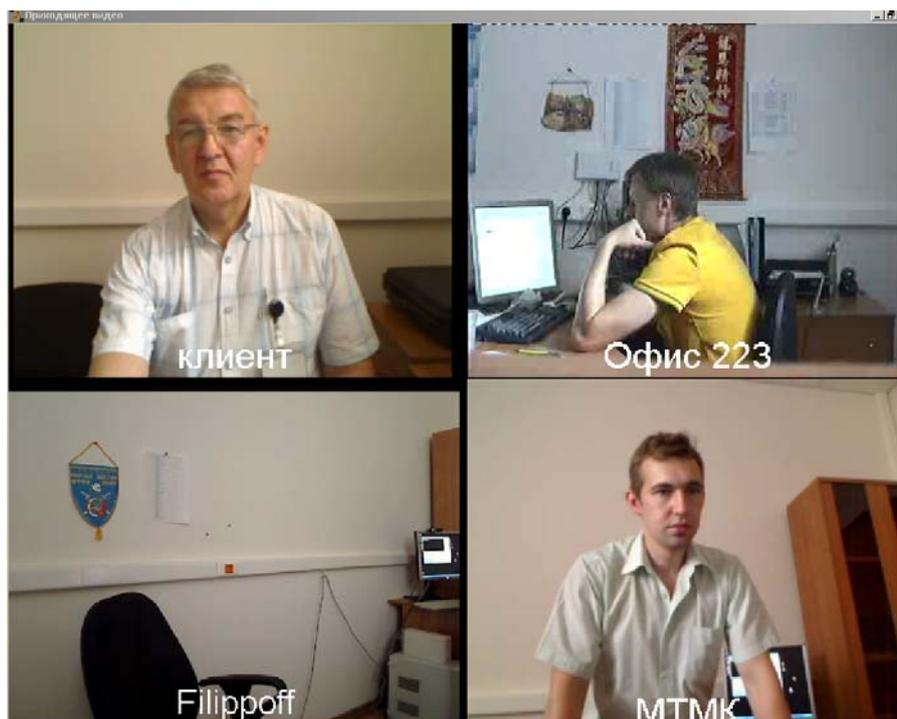


Рис. 1. Экранная форма программы VPhone. Многоточечная ВКС (4 терминала).

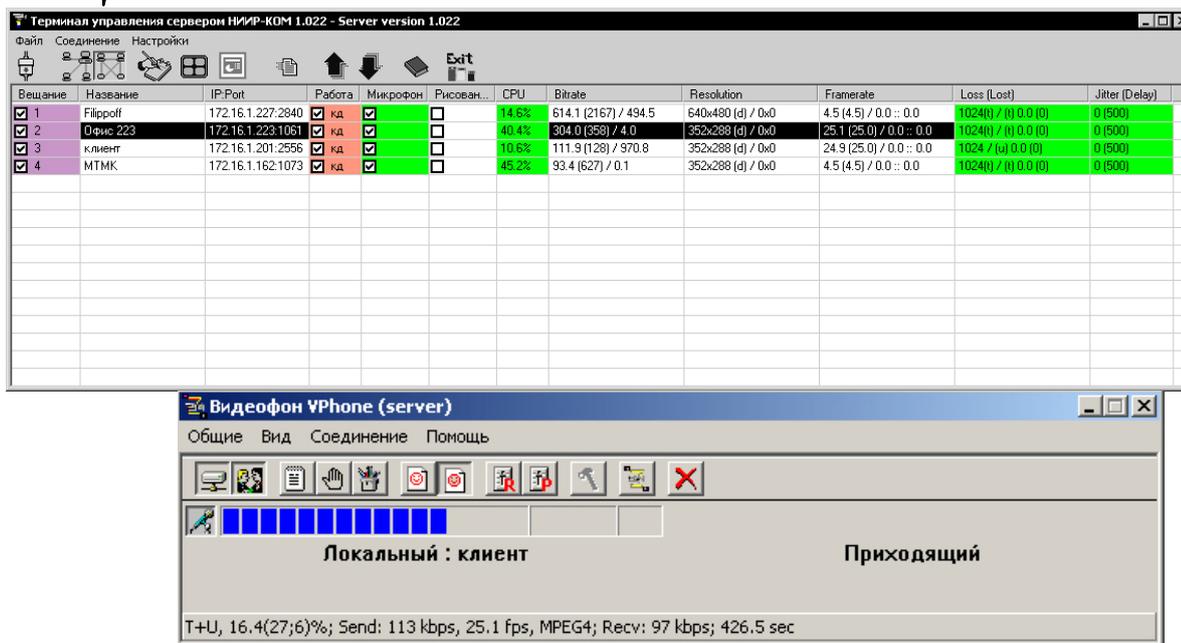


Рис. 2. Экранные формы программы VPhone. Терминалы управления сервером и клиентом.

Дальнейшее распространение комплекса VPhone сдержала его проприетарность. С целью улучшения работы по нестабильным узкополосным IP-сетям VPhone был разработан на собственных протоколах, не совпадающих с H.323, а в стране уже всё шире применялись стандартные системы, с которыми он был несовместим. Кроме этого, коллектив разработчиков покинул НИИ Радио, поскольку его научные планы не нашли поддержки и понимания нового руководства института, а финансирования для продолжения разработки своевременно найдено не было.

В 2007 году коллектив разработчиков, в составе фирмы «НИИР-КОМ», продолжил работу по созданию программных средств видеоконференцсвязи, однако уже не только проприетарных, но и совместимых с рекомендациями H.323. В основу разработок были положены собственные патенты на способы обработки цифровых изображений [35-38]. В 2008 году коллектив составил основу Управления научных исследований Главного радиочастотного центра Роскомнадзора, где и по настоящее время ведёт широкий спектр разработок в области цифрового телевидения и видеоконференцсвязи.

Определяемые стандартами алгоритмы сжатия динамических изображений весьма гибки, что при различных реализациях может обеспечить существенное различие в качестве декодируемой информации при одной и той же степени ее сжатия.

Стандарт видеокодирования H.264/AVC или MPEG-4 Part 10 AVC стандарт постоянно совершенствуется, способен обеспечивать высокое качество воспроизведения динамических изображений при существенно более низких битовых скоростях передачи информации, чем это позволяли стандарты предыдущих поколений.

Среди функциональных возможностей H.264/AVC, которые позволяют поднять эффективность кодирования по отношению к предшествующим стандартам кодирования видео, следует подчеркнуть следующие усовершенствования и новые возможности предсказания содержания кодируемого изображения.

- *Компенсация движения с использованием перемен-*

ных размеров блока, включая малые размеры блока. Минимальный размер блока составляет 4x4 пиксела и позволяет крайне точно выделять области движения.

- *Компенсация движения с точностью до четверти пиксела.* Стандарт повышает точность компенсации движения до четверти пиксела, что особенно актуально для медленного движения. Цветность, как правило, хранится с разрешением, уменьшенным вдвое по вертикали и горизонтали, поэтому компенсация движения для компонента цветности использует точность в одну восьмую пиксела цветности.

- *Вектора движения, выходящие за границы изображения.* Методика экстраполяции за границы изображения, появившаяся как опция в стандарте H.263, включена также в стандарт H.264/AVC.

- *Компенсация движения с несколькими опорными изображениями.* В стандарте H.264/AVC расширяется методика увеличенного выбора опорного изображения, что позволяет кодеру выбирать для компенсации движения между большим количеством изображений, декодированных и сохраненных на декодере. Такое расширение возможностей выбора опорного изображения применяется как для однонаправленного (P-кадры), так и для двунаправленного предсказания (B-кадры).

- *Взвешенное предсказание.* Эта новая возможность стандарта позволяет взвешивать и сдвигать сигнал после компенсации движения на величины, указанные кодером. Такая методика может чрезвычайно сильно поднять эффективность кодирования для сцен с изменением освещенности, а также гибко использоваться для других целей.

- *Направленное пространственное предсказание для внутрикадрового кодирования.* Новая методика экстраполяции краев ранее декодированных частей текущего изображения применяется в областях изображений, кодируемых с помощью внутрикадровой методики (I-кадры). Эта методика повышает качество сигнала, используемого для предсказания, а также позволяет использовать для предсказания соседние области, которые были закодированы не с помощью внутрикадровой методики.

- *Гибкие функции чересстрочного сжатия.* Адаптивное к изображению кодирование полей, позволяющее кодировать каждый кадр как кадр или как пару полей (полукадров) — в зависимости от отсутствия/наличия движения.
- *Квантование.* Логарифмическое управление длиной шага для упрощения распределения цифрового потока кодером и упрощенного вычисления обратной длины квантования. Частотно-оптимизированные матрицы масштабирования квантования, выбираемые кодером для оптимизации квантования на основе человеческих особенностей восприятия.
- *Деблокинговая фильтрация в цикле кодирования.* Кодирование видео, основанное на обработке блоков, приводит к искажениям, называемым «блокинг-эффектом». Причиной может служить как предсказание, так и кодирование остатков предсказания. Применение адаптивного деблокингового фильтра — хорошо известный метод повышения качества восстановленного видео, и при правильном проектировании может повысить как субъективное, так и объективное качество.

В дополнение к улучшенным методам предсказания, другие части стандарта были также расширены с целью повышения эффективности кодирования. Стандарт предусматривает ряд дополнительных функций обработки динамических изображений.

- *Преобразование блоков небольшого размера.* Стандарт основан главным образом на преобразовании блоков 4•4. Это позволяет кодеру представлять сигнал более локально-адаптивным образом, что уменьшает искажения, известные под названием «окантовки». Меньший размер блоков также частично оправдан тем, что позволяет лучше предсказывать содержание видео, но при этом требуется обеспечить области преобразования с границами, соответствующими меньшим областям предсказания.

- *Иерархическое преобразование блоков.* Хотя в большинстве случаев использование малых размеров блоков преобразования (4•4) визуально предпочтительнее, есть некоторые сигналы, содержащие существенную корреляцию, для которых выгоднее использовать более длительные базисные функции. Новый стандарт позволяет делать это двумя способами: (1) с помощью иерархического преобразования (для увеличения эффективного размера блоков) для низкочастотной цветностной информации (8•8); и (2) путем предоставления кодеру возможности выбрать специальный тип внутрикадрового кодирования, позволяющий расширить размер преобразования яркости для низкочастотной информации до размера блока 16•16.

- *Преобразование с использованием 16-битной арифметики.* Все предшествующие стандарты требовали от кодеров и декодеров более сложных вычислений при преобразовании. В то время как ранее требовались, как правило, вычисления с использованием 32-битной арифметики, новый стандарт использует только 16-битную точность вычислений.

- *Точное обратное преобразование.* В предшествующих стандартах видеокодирования преобразование, используемое для представления видео (дискретное косинусное преобразование, ДКП), оговаривалось только в рамках допустимых ошибок из-за практической нереализуемости точного обратного преобразования. В итоге, каждая реализация декодера могла давать слегка различное декодированное видео, обуславливающее несовпадение представ-

ления декодированного видео на кодере и декодере, что приводит к снижению эффективного качества видео. Новый стандарт первым достиг точного совпадения декодированного видео на всех декодерах.

- *Арифметическое энтропийное кодирование.* В стандарт включен передовой метод энтропийного кодирования, известный как арифметическое кодирование. Стандарт использует более эффективную методику — контекстно-адаптивное двоичное арифметическое кодирование CABAC. Предусмотрено также использование контекстно-адаптивное кодирование кодами переменной длины — CAVLC. CAVLC меньшей сложности, чем CABAC, но тем не менее сложнее и эффективнее, чем алгоритмы, применяемые для тех же целей в более ранних технологиях сжатия видео (как правило, это алгоритм Хаффмана).

Программные решения «ДАВ-телекон»

Одним из результатов разработки стал программный терминал ВКС, получивший название «ДАВ-телекон» (Данные-Аудио-Видео-телеконференция), который представляет собой эффективное средство видеосвязи, обладающее дружественным программным графическим интерфейсом (ПГИ) с развитыми возможностями управления параметрами текущего сеанса и адресной книгой (рис. 3-7).

Программа «ДАВ-телекон» зарегистрирована 21-го января 2010 г. в государственном Реестре программ для ЭВМ за № 2010611864.

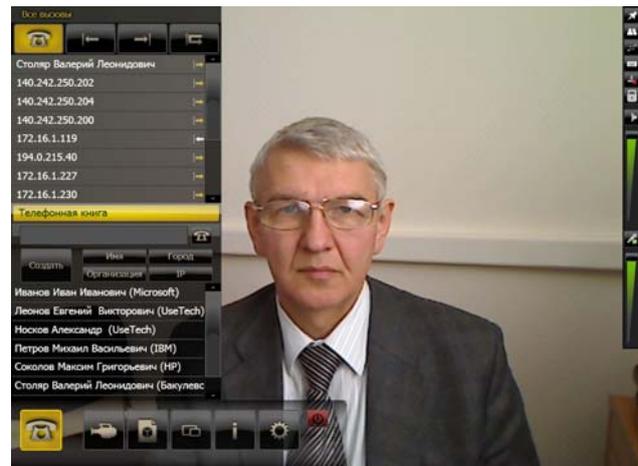


Рис. 3. ПГИ – меню и интерфейс вызовов абонентов.

На рис. 3 представлена исходная форма ПГИ, предназначенная для вызова удалённых абонентов по записи в адресной книге, IP-адресу или списку последних соединений. Кнопки внизу предназначены для вызова основных функций меню: настройки, статистики параметров текущего соединения, управления выводом входящих видеопотоков, управления запуском приложений, назначения источников исходящих видеопотоков, управления соединениями. Кнопки и движки на правой вертикальной панели предназначены для: управления громкостью входящего и исходящего звука и его блокирования, вызова виртуальной клавиатуры, переключения из полноэкранный режим в оконный и наоборот, включения/отключения исходящего видеопотока, включения/отключения режима двух видеопотоков, включения/отключения режима «поверх всех окон».

На рис. 4 - 7 показаны окна настройки параметров конкретных функций и окно редактирования карточки абонента.

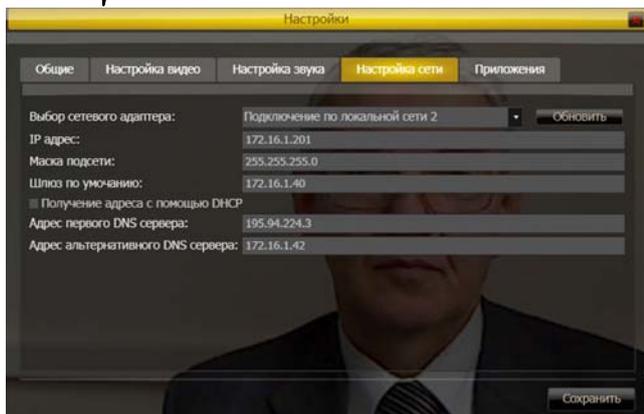


Рис. 4. Окно настроек параметров сети.

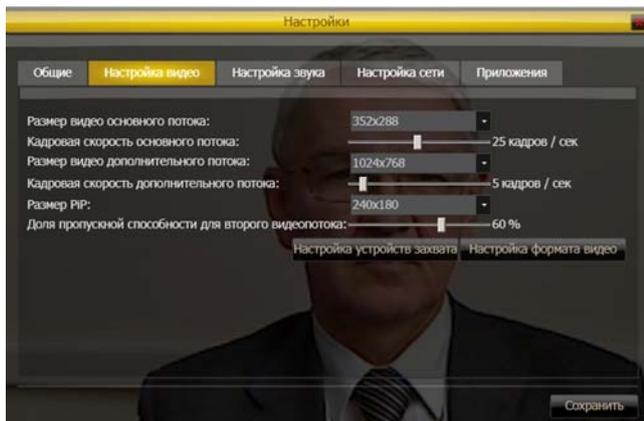


Рис. 5. Окно настроек параметров видеопотоков.

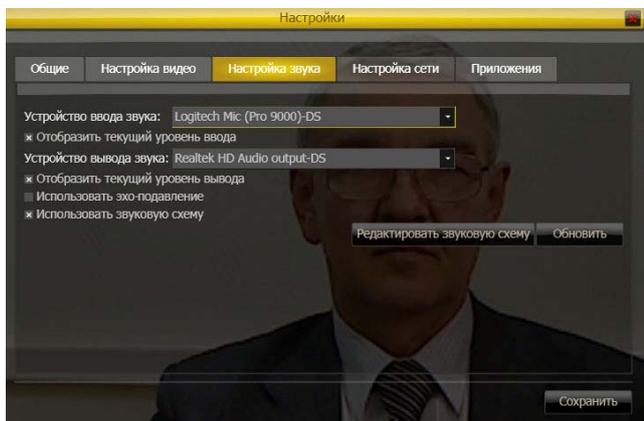


Рис. 6. Окно настроек параметров аудиоподсистемы.



Рис. 7. Режим редактирования карточки абонента.

В настоящее время программный терминал видеоконференцсвязи «ДАВ-телекон», работающий в IP-сетях в соответствии с рекомендациями H.323 Международного Союза Электросвязи, в чём-то несколько уступает по функциональности аппаратным комплексам ведущих производителей, в чём-то превосходит их, однако имеющихся возможностей вполне достаточно для подавляющего большинства пользователей.

Поскольку различия между программами «ДАВ-телекон» и «VPhone» касаются, прежде всего, стека протоколов, организующих взаимодействие терминалов между собой и терминалов с сервером, очевидно, что ПГИ может быть перенесён в программу «VPhone», что существенно упростит её применение неподготовленными пользователями. И оба программных продукта, имеющих сходный интерфейс, могут применяться в соответствии со своими особенностями: «VPhone» – в тех случаях, когда требуется устойчивая работа на узкополосных нестабильных сетях и/или проведение многоточечных видеоконференций без дорогостоящих аппаратных видеосерверов, «ДАВ-телекон» – когда нужен бюджетный терминал, совместимый с H.323.

Подводя итоги

Исходя из всего вышесказанного разработчики современных средств видеоконференцсвязи должны ориентироваться, прежде всего, на актуальные потребности наиболее массового потребителя, как реального, так и потенциального. А значит, придётся признать, что будущее скорее за программными решениями, чем за аппаратными.

На выставке Связь-Экспокомм-2010 в мае 2010 года пермской компанией «Интеллектуальные системы и технологии» были представлены серийные видеотаксофоны. В июле появилась информация о начале их внедрения в Тульской области [39]. Однако и в их основе лежат не специализированные процессоры (аппаратное решение), а компьютеры общего назначения и программное обеспечение видеоконференцсвязи.

У каждого, кто внимательно знакомится с реальными потребностями и предложениями в области видеоконференцсвязи, возникает вопрос: а насколько оправдано стремление к совместимости со стандартом H.323, разработанным исходно для более широкого класса телекоммуникационных услуг и обладающим избыточной сложностью для реализации большинства запросов пользователей ВКС? Не слишком ли высока плата за невостребованную универсальность? Может быть, следует в качестве российского стандарта принять существующую «де-факто» разработку, позволяющую с приемлемым качеством работать по узкополосным и нестабильным сетям связи, а в широкополосных сетях обеспечивающую качество сравнимое с дорогостоящей стандартизированной аппаратурой?

Что касается тех немногих ситуаций, когда действительно требуется полная совместимость с аппаратурой ведущих мировых производителей [2, 3], то проблема относительно просто решается путём применения программного шлюза, обеспечивающего перекодирование стека протоколов.

Сегодня технологиями ВКС пользуется относительно небольшой процент населения, а завтра они станут

массовыми. И тогда вопрос о соотношении цена/качество выйдет на первый план. Российские разработчики готовы предложить потребителю необходимые программные средства по разумным ценам. Вопрос только за нормативным обеспечением и решением руководства телекоммуникационной отрасли.

Источники информации:

1. http://www.rol.ru/it/news/96/12/03_22.htm (04.03.2010).
2. <http://www.polycom.com>, <http://www.sony.ru>,
<http://www.emblaze-vcon.ru>, <http://www.vtel.com>,
<http://www.aethra.ru> (04.03.2010).
3. <http://www.tandbergussia.ru> (04.03.2010).
4. <http://www.osp.ru/text/print/302/179460.html> (04.03.2010).
5. http://obsud.sah.sudrf.ru/modules.php?name=press_dep&op=1&did=283 (04.03.2010).
6. http://www.muzbazar.ru/Biblioteka/Videokonferencii_i_ih_ispolzovanie.html (04.03.2010).
7. <http://www.videosvyaz.ru> (04.03.2010).
8. http://www.tandbergussia.ru/ind_focus/energy/index.jsp (04.03.2010).
9. <http://www.pentagonus.ru/publ/3-1-0-669> (05.03.2010).
10. <http://www.connect.ru/article.asp?id=9552> (05.03.2010).
11. <http://www.pcweek.ru/themes/detail.php?ID=82777> (05.03.2010).
12. <http://sanweras.pp.ua/videokonferentsyvyaz-2/vpervie-v-rossii-bil-proveden-dopros-svidetelya-s-ispolzovaniem-videokonferentsyvyazi-639.html> (05.03.2010)
13. <http://www.pfo.ru/?id=19837> (05.03.2010).
14. <http://www.kremlin.ru/news/6148> (05.03.2010).
15. <http://www.rian.ru/politics/20050610/40504925.html> (06.03.2010).
16. <http://tele.med.ru/u1.htm> (06.03.2010).
17. <http://www.telemed.ru> (06.03.2010).
18. http://www.pmc.ru/Products_and_Services/Telemedicine/GRCCTelemed (09.03.2010).
19. http://www.rosnou.ru/important/filial_konf (09.03.2010).
20. http://www.vfmgiu.ru/chto_novogo_v_visshem_obr_zovanii_450/videokonferencyvyaz_prob-lemu_reshenie_466/index.html (09.03.2010).
21. <http://www.connect.ru/newsdetail.asp?id=6543> (09.03.2010).
22. Фёдоров В.Ф. Российские услуги связи с точки зрения телемедицины. В сб. научно-практической конф. «Мобильные телемедицинские комплексы. Домашняя телемедицина». Ростов-на-Дону, 2005 г. С. 75-82.
23. <http://www.cnews.ru/reviews/index.shtml?2008/05/19/300715> (09.03.2010)..
24. <http://www.bs-media.ru/catalogcat/video/1147247225~1148886071> (09.03.2010).
25. <http://www.tandbergussia.ru/products/pc-video-conferencing.jsp> (09.03.2010).
26. http://www.emblaze-vcon.ru/solut_nastol.htm (09.03.2010).
27. <http://office.microsoft.com/en-us/livemeeting/HA102403231033.aspx> (10.03.2010).
28. <http://ru.wikipedia.org/wiki/Skype> (10.03.2010).
29. <http://www.videoport.ru> (10.03.2010).
30. http://www.logitech.com/index.cfm/webcam_communication/s&cl=ru,ru (10.03.2010).
31. <http://telecom.cnews.ru/reviews/index.shtml?2010/03/19/383377> (10.03.2010).
32. <http://www.telesputnik.ru/archive/64/article/40.html> (11.03.2010).
33. http://www.muh.ru/stud_practice.htm (11.03.2010).
34. <http://steelsoft.ru/rosnou/> (11.03.2010)
35. Дворкович В.П., Мохин Г.Н., Нечепав В.В., Дворкович А.В. Патент 2122295 (РФ). Способ покадрового сжатия изображений.// Б.И. - 1998.
36. Дворкович А.В., Дворкович В.П., Зубарев Ю.Б., Соколов А.Ю. Патент 2137194 (РФ). Способ анализа векторов движения деталей в динамических изображениях.// Б.И. - 1999. 37. Дворкович А.В., Дворкович В.П., Соколов А.Ю. Патент 2182727 (РФ). Способ поиска векторов движения деталей в динамических изображениях.// Б.И. - 2002.
37. Дворкович А.В., Дворкович В.П., Мохин Г.Н., Соколов А.Ю. Патент 2182746 (РФ). Способ цифровой обработки динамических изображений.// Б.И. - 2002.
39. <http://telecom.cnews.ru/news/top/index.shtml?2010/07/26/402531> (26.07.2010).

Уважаемые коллеги!

Для тех, кто не успел оформить подписку на первое полугодие 2011 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

РОССИЙСКИЕ ИННОВАЦИОННЫЕ ТЕХНОЛОГИИ В ПРИЕМНОЙ АППАРАТУРЕ ЦИФРОВОГО ТЕЛЕВИДЕНИЯ НА 22-ой МЕЖДУНАРОДНОЙ ВЫСТАВКЕ "СВЯЗЬ-ЭКСПОКОММ-2010"

Вилкова Н.Н., Генеральный директор ЗАО "МНИТИ", к.т.н.

Быструшкин К.Н., заместитель Генерального директора по научной работе ЗАО «МНИТИ», к.т.н.

В период с 11 по 14 мая 2010 г. в ЦВК Экспоцентр "Красная Пресня" проходила 22-я Международная выставка "СВЯЗЬ-ЭКСПОКОММ-2010"/"ТВЧ Россия 2010". Совместная экспозиция "Цифровое телевидение" Минкомсвязи РФ (ФГУП "РТРС"), Минпромторга РФ (ЗАО "МНИТИ" и Ассоциация "АРПАТ") и ФГУП ВГТРК была развернута в павильоне 2 зал 3 на стенде № 23D70 общей площадью 432 м². Экспозиция "Цифровое телевидение" наглядно показала реализацию ФЦП "Развитие цифрового телерадиовещания в РФ на 2009-2015 годы", а также перспективы развития технологий телерадиовещания в России до 2020 года.

Основные цели совместной экспозиции:

1. Показ первых итогов реализации ФЦП "Развитие телерадиовещания в РФ на 2009-2015 годы" – комплексный стенд "Сделано в России" Ассоциации "Арпат" – демонстрация готовности отечественной промышленности принять активное участие в реализации ФЦП в части создания материально-технической базы приемной и передающей сети цифрового телевидения.

2. Презентация ФГУП ВГТРК – Парад Победы 9 мая 2010 года на светодиодном LED экране Full HDTV размером 3x8 м (производство ООО "Дисплейные системы") и в варианте 3D на стереоскопических очковых и безочковых телевизорах (совместно с ЗАО "МНИТИ", ЗАО "Элекард", ЗАО "Триаксес Вижн", ЗАО "Тринити Солюшнс" и компанией "JC System Integration").

3. Презентация концепции развития технологий цифрового телевидения в России до 2020 г. ("Дорожная карта") – ЗАО "МНИТИ", ФГУП РТРС и ФГУП ВГТРК.

4. Показ новейших телевизионных технологий: стереоскопического 3D телевидения, интерактивного телевидения, в том числе возможность подключения к информационной системе "Электронное правительство РФ", мобильного телевидения DVB-SH, второго поколения наземного цифрового телевидения DVB-T2, сверхскоростной телевизионной камеры 2K x 2000 кадров.

Совместная экспозиция "Цифровое телевидение" и ее составные части ("Сделано в России": "Приемная аппаратура цифрового телевидения 2010", Ассоциация "АРПАТ" - Минпромторг РФ, "Российские 3D технологии", "Гибридные технологии цифрового телевидения" и др.) должна была наглядно показать, что в России имеется единая долгосрочная концепция развития телевизионных технологий, совместная реализация которой начинается уже сегодня. Сама экспозиция, которая отразила наиболее значимые мировые тренды развития цифрового телевидения, показала, что социальный заказ со стороны руководства РФ на инновационное развитие прорывных технологий и создание в России опережающего научного и технического

Российская экспозиция "Цифровое телевидение", на которой были представлены широкие модельные ряды приемной и передающей аппаратуры, показала, что отечественная промышленность готова принять активное участие в реализации ФЦП "Развитие телерадиовещания в РФ на 2009 -2015 годы". Экспозиция отразила также наиболее значимые мировые тренды развития цифрового телевидения, в том числе аппаратуру 3D телевидения и телевидения второго поколения DVB-T2.

задела уже реализуется.

Для достижения указанных выше целей экспозиция при внешнем стилистическом единстве состояла из нескольких функциональных зон:

1. **Российские дисплейные технологии** (ВГТРК и LED экран холдинга "Инкотекс" (ООО "Дисплейные системы").

2. **Гибридные технологии цифрового телевидения** (интерактивное и IPTV (ОАО "Уфимский завод "Промсвязь", ОАО "Ставропольский завод "Сигнал") + экспозиция Корпорации "General Satellite").

3. **Второе поколение цифрового телевидения** (DVB-SH + DVB-T2 – ФГУП РТРС, компания ALCATEL, ЗАО "МНИТИ" и ОАО "Ставропольский завод "Сигнал").

4. **Спутниковое 3D телевидение** (Корпорация "General Satellite").

5. **3D развлечения и досуг** (ООО "Самсунг Электроникс Рус компании Калуга").

6. **3D в реальном времени** (ЗАО "МНИТИ", ЗАО "Триаксес Вижн", ЗАО "Элекард", ЗАО "Тринити Солюшнс", JC System Integration).

7. **"Сделано в России" – Приемная аппаратура цифрового телевидения 2010** (ЗАО "Интеркросс" (г. Рязань), ОАО "Радиозавод" (г. Пенза), ООО "Вестел трейд" (г. Александров), ООО "Самсунг Электроникс Рус компании Калуга"; ОАО "Уфимский завод "Промсвязь" (г. Уфа), ОАО "Ставропольский завод "Сигнал", Корпорация "General Satellite").

8. **"Сделано в России" – Ассоциация "АРПАТ"** – комплексное решение сетей цифрового телевидения (группа компаний г. Санкт-Петербурга (ОАО "МАРТ", ОАО "РИМР", ЗАО "Завод им. Козицкого") + ФГУП "ОмПО "Иртыш" (г. Омск)).

9. **"Сделано в России" – "Коллективный прием цифрового телевидения"** - УКПТ (Ц/А) 2-го поколения – ЗАО "МНИТИ", ООО "Телемак".

10. **Российские 3D технологии – 3D телестудия** (ЗАО "МНИТИ", ООО "Триаксес Вижн" (г. Томск), ЗАО "Элекард" (г. Томск), JC System Integration (г. Москва).

11. **Российские 3D технологии – видеомонтаж 3D** (ЗАО "МНИТИ" ЗАО "Триаксес Вижн" (г. Томск), ЗАО "Элекард" (г. Томск), JC System Integration (г.Москва).

12. **Телевидение XXI века** (2K телекамера – ФГУП ВГТРК).



Рис.1. Совместная экспозиция ФГУП РТРС, ЗАО «МНИТИ» и ФГУП ВГТРК на выставке "Связь-Экспокомм 2010"

Российские технологии 3D цифрового телевизионного вещания

ЗАО «МНИТИ»; ООО «Трижес Викж»; ЗАО «Элекрад»; ЗАО «Тринити Солюшнс»

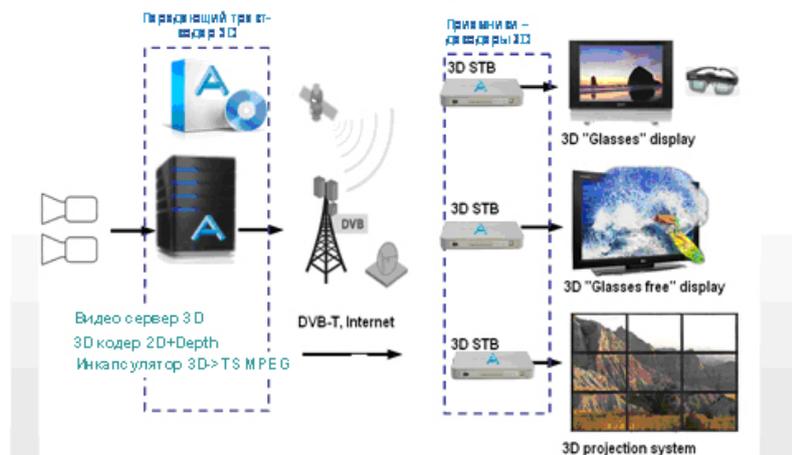
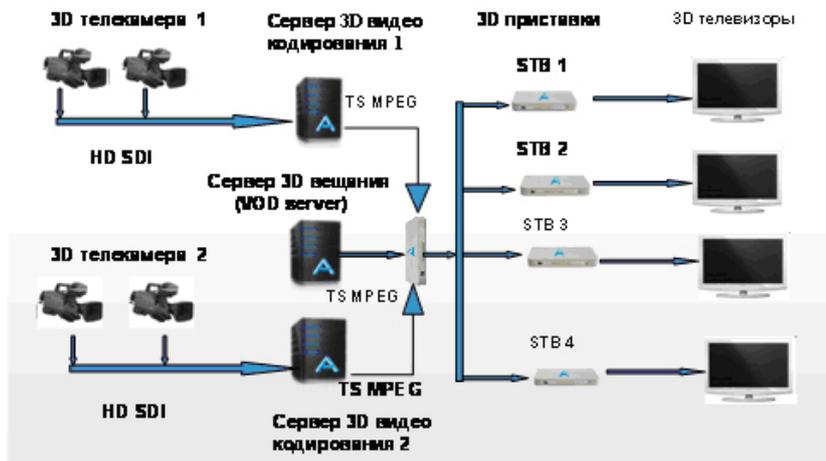
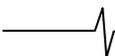


Схема 3D цифрового телевизионного вещания в реальном масштабе времени на выставке «Связь-Экспокомм 2010»

ЗАО «МНИТИ»; ООО «Трижес Викж»; ЗАО «Элекрад»; JSC System Integration; ЗАО «Трижес Солюшнс»





На стенде "Цифровое телевидение" ЗАО "МНИТИ" и его партнеры впервые показали в режиме реального времени трансляцию стереоскопического изображения с 3D телевизионных камер на безочковые стереотелевизоры. Кроме того, на очковых и безочковых 3D дисплеях был показан в режиме 3D Парад Победы, снятый операторами ВГТРК 9 мая 2010 года на Красной Площади. Эта демонстрация вызвала большой интерес посетителей выставки и свидетельствует о высоком уровне российских разработчиков.

Среди других достижений российских компаний необходимо отметить первую демонстрацию российского прототипа цифровой приставки для телевидения второго поколения DVB-T2, который представили ОАО "Ставропольский завод "Сигнал" и ЗАО "МНИТИ". Как известно, в Европе развертывание опытных сетей цифрового вещания DVB-T2 началось только в прошлом году, и до сих пор ведущие мировые компании не начали серийного производства таких приемников, ввиду сложности задачи их разработки.

Еще одним наглядным подтверждением высокого уровня российских инновационных технологий является разработка целым рядом российских компаний цифровых приемников нового поколения со встроенным Internet-браузером. Подобная цифровая приставка, подключенная к сети Internet, позволяет зрителю при помощи подключенного к приставке телевизора получать доступ к полному спектру информационных услуг, доступным сегодня только пользователям персональных компьютеров. На стенде "Цифровое телевидение" ОАО "Уфимский завод "Промсвязь" успешно демонстрировало работу такой цифровой приставки собственной разработки, а также созданного на ее основе Интерактивного информационного терминала, подготавливаемого к серийному производству в рамках программы "Электронное правительство".

Компания ООО "Самсунг Электроник Рус Калуга" и НПФ "Эра" показали действующий образец Internet-телевизора со встроенным браузером для навигации по "Всемирной паутине". Таким образом, этот год станет переломным в окончательной конвергенции компьютерных и телевизионных технологий на технологической платформе цифрового телевидения, о наступлении которой специалисты говорят последние 15-20 лет.

В области приемной аппаратуры цифрового телевидения большой интерес посетителей вызвали цифровые приставки и цифровые телевизоры из модельных рядов 2010 года. От моделей предыдущего поколения они отличаются возросшей вычислительной мощностью используемых в схеме приемников цифровых процессоров, что позволило превратить цифровые приемники в настоящие мультимедийные центры. Большинство моделей 2010 года помимо приема программ цифрового телевидения DVB-T/MPEG-4 H/264 AVC позволяют считывать информацию с подключенных к ним флеш-карт (например, фотографии с цифровых фотоаппаратов), подключать к ним по USB порту внешние накопители HDD (при этом практически бесплатно приставка или телевизор превращаются в цифровой видеомаягнитофон) и даже выход в сеть Internet в наиболее продвинутых моделях.

Большой интерес посетителей вызвала демонстрация второго поколения устройства коллективного приема УКПТ

(Ц/А), разработанного ЗАО "МНИТИ" и ООО "Телемак" (г. Саратов). С помощью этого устройства, преобразующего программы эфирного цифрового телевидения DVB-T/MPEG-4 H/264 AVC в аналоговые телевизионные сигналы SECAM D/K, телезрители имеют возможность принимать цифровые телепрограммы на имеющиеся у них аналоговые телевизоры. Так как устройства УКПТ (Ц/А) позволяют подключить к их выходу до 300 квартир, стоимость "цифровизации" одной квартиры при использовании коллективного приема составляет всего 300-400 руб., тогда как цена самой простой цифровой приставки сегодня не ниже 2400 руб. Экономия средств и удобство для телезрителей этого варианта приобщения к цифровому телевидению настолько очевидны, что на сегодняшний день из различных регионов России начали поступать заявки на приобретение в опытную эксплуатацию устройств коллективного приема. В настоящий момент они проходят испытания в Казани, Зеленодольске (Республика Татарстан), Твери, Санкт-Петербурге и Курске. Один образец УКПТ (Ц/А) к Дню Победы был передан в дар Курскому военному госпиталю.

В этом году ЗАО "МНИТИ" и ООО "Телемак" провели радикальную модернизацию устройства УКПТ (Ц/А), создав практически новое поколение этой аппаратуры. Во втором поколении почти вдвое уменьшены габариты и масса устройства, введен дистанционный мониторинг их рабочих режимов по сети GSM, а также добавлена функция распространения программ цифрового радио по сетям коллективного приема в аналоговом формате в FM диапазоне.

Высокий уровень показанных на экспозиции "Цифровое телевидение" технологий привлек к ней повышенное внимание посетителей и руководителей различного уровня. За время работы выставки экспозицию посетили Вице-Премьер Правительства РФ С.С. Собянин, Министр связи и массовых коммуникаций России И.О. Щеголев и его заместители – Н.С. Мардер, А.А. Жаров, Д.С. Северов, Заместитель Министра промышленности и торговли России Ю.И. Борисов, Директор департамента радиозлектронной промышленности Минпромторга РФ В.Н. Минаев и его заместитель А.Е. Суворов, Директор департамента новых технологий в телерадиовещании Минкомсвязи России В.П. Стыцко и еще целый ряд высокопоставленных руководителей. Работа выставки широко освещалась в средствах массовой информации и, в частности, на федеральных телеканалах.

Таким образом, успешное участие российских компаний в выставке «Связь-Экспокомм 2010» наглядно показало наличие в РФ опережающего научно-технического задела и высокого научного и технического потенциала российской науки и техники в области цифрового телерадиовещания. Благодаря координирующей роли Министерства промышленности и торговли России и его головной научной организации в области цифрового телевидения ЗАО "МНИТИ", в стране создана научная и производственная кооперация, позволяющая отечественной промышленности активно включиться в борьбу за огромный рынок приемной и передающей аппаратуры, формирующийся при переходе от аналогового к цифровому телевизионному вещанию. И в этом заключается один из важнейших результатов прошедшей выставки "Связь-Экспокомм 2010".

НОВАЯ РУБРИКА: «ПЕРВЫЕ ШАГИ В НАУКЕ»

УДК 004.932

ВОЗМОЖНОСТИ ИСПОЛЬЗОВАНИЯ ФАКТОРА НАСЛЕДСТВЕННОСТИ ПРИЗНАКОВ НА ПОСЛЕДОВАТЕЛЬНОСТИ ИЗОБРАЖЕНИЙ ДЛЯ ИХ АНАЛИЗА

Зуев В.В., аспирант кафедры САПР ВС, Муромский институт (филиал) Владимирского государственного университета, zuev_v.lan@mit.ru, Zuev-V-V@yandex.ru

Жизнякав А.Л., научный руководитель, зав. кафедрой САПР ЭС, Муромский институт (филиал) Владимирского государственного университета, д.т.н.

Ключевые слова: контуры, изображения, размерность, наследственность признаков, эквализация,

Актуальность темы

Зрительные образы являются для человека основой восприятия окружающего мира. Поэтому во многих отраслях науки и техники изображения являются основным источником информации об исследуемом объекте или явлении. На сегодняшний день, в связи с постоянно растущим потоком этой информации и повышением требований к точности и оперативности извлечения необходимых данных, обработка разного рода изображений немислима без использования средств вычислительной техники.

При анализе изображений одной из основных проблем является значительный объем входных данных. Решением этой проблемы является обработка не всех пикселей изображения, а наиболее информативной части. Такой частью являются, например, контуры. Контурный анализ представляет собой совокупность методов выделения, преобразования и описания контуров изображений. Контуры изображений рассматриваются как k -мерные векторы, т.е. вектор-контур. Для того, чтобы обработать контур на ЭВМ, его необходимо закодировать. В [1] рассматриваются некоторые виды кодирования контуров, а именно: а) код Фримена, б) двумерный код, в) полигональное представление, г) комплексно-значный код. Наиболее универсальным для сравнения контуров считается пространство комплексных векторов C^k . Это пространство инвариантно к сдвигу, повороту и изменению масштаба контура. Для сравнения контуров в пространстве предлагается использовать свойство нормированного скалярного произведения (НСП), единичное значение которого обозначает идентичность сравниваемых контуров, а нулевое их полное несоответствие.

Для этого вычисляют взаимнокорреляционную функцию (ВКФ) двух анализируемых контуров, и максимальное значение полученной функции принимают за меру близости. Операция вычисления ВКФ является трудоемкой, и для пространства C^k количество обращений к элементам вектора составляет k^2 .

В данной статье предлагается подход, позволяющий значительно сократить количество обращений к элементам контура, а значит и трудоемкость алгоритма вычисления ВКФ.

На основании экспериментальных исследований процессов, протекающих при эквализации контура к новой размерности, предлагается подход к обработке изображений, основанный на использовании фактора наследственности контура при его эквализации к различным размерностям.

Выявление фактора наследственности

Эквализацией называется процедура выравнивания размерности двух вектор-контуров [1]. Причем дополнительных условий на способ выравнивания эквализация не накладывает. Основным моментом является одинаковая размерность получаемых вектор-контуров. Обычно, после операции выделения, вектор-контур состоит из элементарных векторов. Элементарные векторы соединяют центры соседних пикселей, принадлежащих линии контура. Часто такое представление контура является избыточным, и контур можно описать ВК меньшей размерности без потери информации или с незначительной потерей.

Пусть требуется найти меру близости для двух контуров размерностей k_1 и k_2 для различных размерностей. Проведем эквализацию контуров в пространства размерностями $L = \{3, \dots, \max(k_1, k_2)\}$. Вычислим для каждой размерности пространства ВКФ. На рис.1 представлен график значения меры близости пары контуров тестовых изображений в зависимости от размерности пространства L и смещения начального элементарного вектора. Явно видно, что общий вид ВКФ при различных размерностях пространства имеет большую степень подобия.

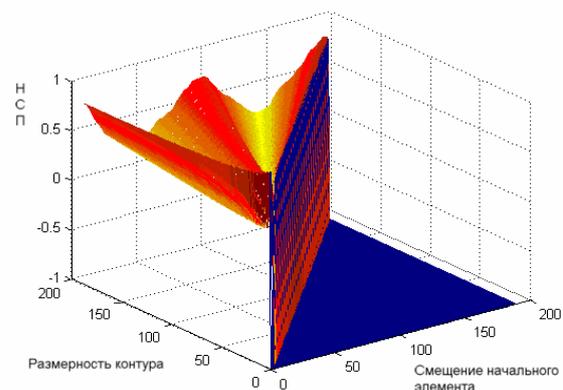
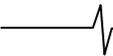


Рис. 1. График «Размерность контура, смещение индекса, значение НСП» для двух контуров. Значение «-1» присваивалось для значения НСП при смещении начального элемента контура, превышающем мерность пространства.



В [2] описана теория, позволяющая обосновать полученное подобие. Эквализация подразумевает преобразование размерности контура в один этап, т.е. из контура размерности k получается контур размерности m . Обозначим это преобразование как некоторый оператор $T_{k,m}$.

Сам процесс можно проводить и поэтапно, то есть проводить преобразование вектора не к размерности m , а к некоторой промежуточной размерности m' такой, что:

$$m' \in \begin{cases} \{m+1, \dots, k-1\}, & k > m \\ \{k+1, \dots, m-1\}, & k < m \end{cases}$$

где m' , m , k - размерности векторов.

При таком подходе преобразование $T_{k,m}$ изображения контура V^k в контур V^m можно представить в виде некоторой последовательности:

$$V^m = \text{last} \left(\left\{ T_{k,m_1} V^k, T_{m_1,m_2} V^{m_1}, \dots, T_{m_{i-1},m_i} V^{m_{i-1}}, T_{m_i,m} V^{m_i} \right\} \right).$$

Здесь: T - оператор преобразования, $k, m_1, m_2, m(i-1), m_i, m$ - размерности векторов, i - некоторый индекс, last - функция выбора последнего элемента последовательности.

В [2] указывается, что в таком случае можно говорить о наличии фактора наследственности. Наследственность признака выражается в том, что на подпоследовательности изображений содержится один и тот же признак с различной степенью отчетливости. На рис. 1 фактор наследственности выражен в гомотетии ВКФ для различных пространств.

Исследование изменения значения меры близости при эквализации ВК

Поскольку сравнение контуров основано на вычислении их меры близости, то необходимо проанализировать её изменение при эквализации контуров в пространства различной размерности.

В результате экспериментальной проверки было получено:

1) Мера близости контура с самим собой равна 1 при любой размерности пространства. Это очевидный результат, поскольку результат эквализации контура один и тот же.

2) Мера близости двух различных контуров при уменьшении их размерности имеет тенденцию к увеличению с незначительными осцилляциями, что отображено на рис. 2. Но такая зависимость может нарушаться локальными уменьшениями различной степени (рис. 3а) или значительными осцилляциями, как показано на рис. 3б.

3) Для определения наличия влияния начального элемента контура было вычислено значение автокорреляционной функции $AK\Phi(25)$ для вектор-контура изображения. Результат показан на рис. 4. Это единственный случай, когда наблюдается общее уменьшение меры сходства.

Во всех случаях при уменьшении размерности пространства в 2 раза значительного изменения меры близости не происходит. Как показали расчеты, максимальное абсолютное отклонение составляет 0,0153, а относительное 0,0196.

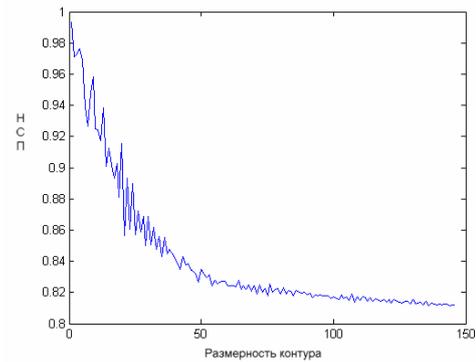
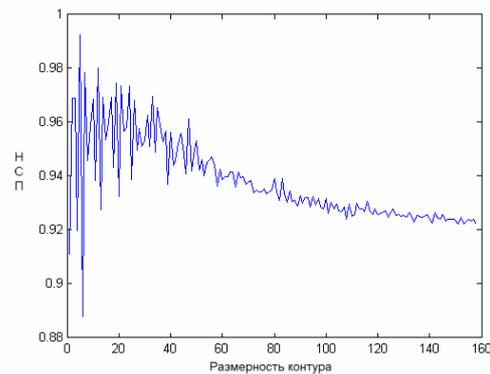
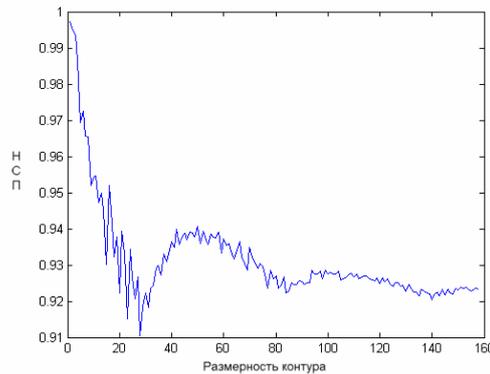


Рис. 2. Зависимость меры близости двух контуров от размерности пространства.



а) значительные осцилляции значения меры близости



б) локальные уменьшения значения меры близости

Рис. 3. Зависимость меры близости двух контуров от размерности пространства. По оси X отложена размерность контуров, по Y мера близости.

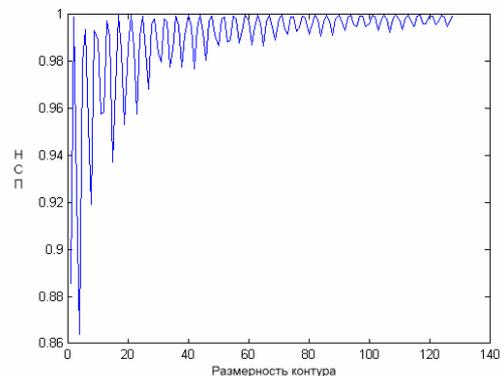


Рис. 4. Значение меры близости двух одинаковых контуров с различным начальным элементарным вектором. По оси X отложена размерность контуров, по Y мера близости вектор-контуров.

Стоит отметить, что при эквализации контура происходят незначительные искажения, которые, постепенно накапливаясь, могут изменять значение меры близости контуров, поэтому необходимо уметь оценивать её изменение.

Исследование влияния количества последовательных эквализаций вектор-контуров на значение их меры близости

Как было сказано выше, эквализацию ВК можно проводить не в один, а в несколько этапов. В литературе нет исследований по этому вопросу. Необходимость исследования этого вопроса можно проиллюстрировать на следующем примере. Пусть на изображении дано n ВК различной размерности. Необходимо сравнить попарно все представленные контуры. Первый способ заключается в том, что все пары контуров сравниваются независимо, то есть для некоторого фиксированного вектора V (для определённости с максимальной размерностью m) необходимо провести $n-1$ эквализаций. Соответственно количество обращений к элементам контура будет равно $(n-1)*m$. Другой способ работы заключается в том, что для каждой пары векторов эквализировать можно не исходный вектор V , а вектор V_i , полученный эквализацией вектора V к меньшей размерности. Тогда количество обращений к элементам вектора сократится и будет меньше $(n-1)*m$. При этом, однако, неизвестно, как будут соотноситься значения меры близости контуров, полученные первым и вторым способом. Возникает необходимость выбора, какой вариант работы предпочесть? Данное исследование призвано ответить на этот вопрос. Исследование проводилось по следующей схеме. Для двух контуров вычислялась мера близости из размерности n к размерности m , получаемая последовательной эквализацией с шагом в 1,2,3,... $n-m$ размерности. Результаты исследований представлены на рис. 5.

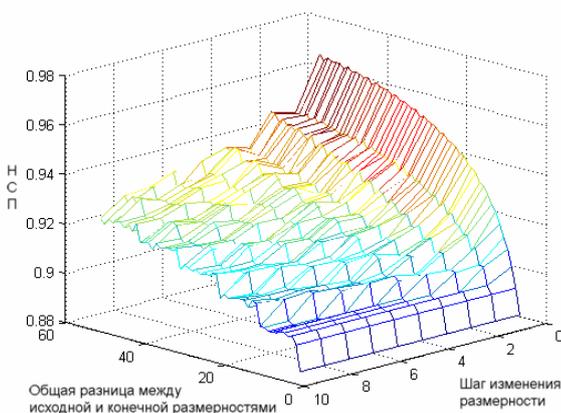


Рис. 5. График влияния количества последовательных эквализаций контуров на значение их меры близости.

Из рис. 5 видна общая зависимость, которую можно сформулировать следующим образом: чем больше промежуточных эквализаций контуров, тем меньше точность значения меры близости контуров.

Поэтому, если аппаратных ресурсов достаточно, то рекомендуется первый способ сравнения множества контуров, иначе при втором способе работы необходимо учитывать, что полученная мера близости будет иметь завышенное значение.

Исследование изменения энергии контура при изменении размерности пространства

Важной характеристикой вектор-контура является его энергия. На рис. 6 показано изменение значения энергии вектор-контура при эквализации к различным размерностям, с учетом разного начального элементарного вектора.

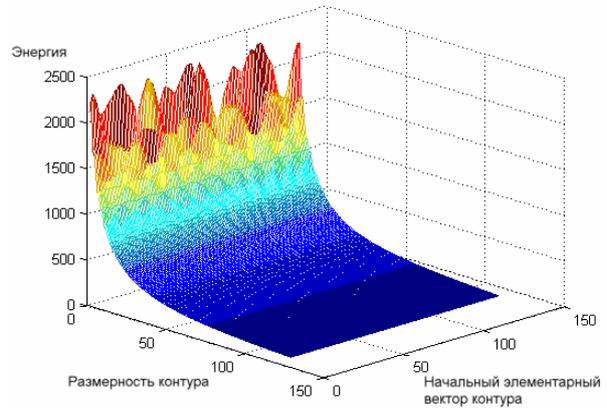


Рис. 6. Изменение значения энергии вектор-контура, при его эквализации к различным размерностям, с учетом разного начального элементарного вектора.

Из рис.6 видно, что при уменьшении размерности пространства модуль вектора контура увеличивается, но в значительной мере зависит от начального элементарного вектора при малых значениях размерности.

Пусть имеются вектор-контур V_1 размерности m , у которого длины всех элементарных векторов равны приблизительно b , и вектор-контур V_2 , полученный эквализацией вектора V_1 к размерности n . Предположим, что длины элементарных векторов V_2 приблизительно равны, а также равны длины контуров V_1 и V_2 . Пусть $m/n = a$. Вычислим отношение энергии вектор-контура V_2 к энергии вектор-контура V_1

$$\frac{E_2}{E_1} = \frac{b^2 \cdot m}{b^2 \cdot m^2 / n} = \frac{n}{m} = \frac{1}{a}.$$

Таким образом, изменение энергии контура при эквализации обратно пропорционально изменению размерности вектор-контура. Результат экспериментальной проверки представлен на рис.7.

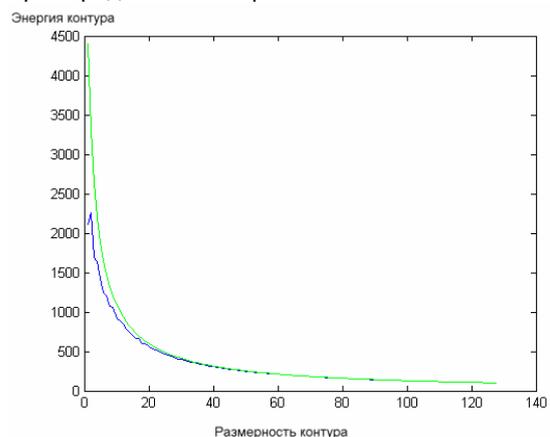
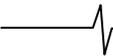


Рис.7. Графики изменения значения энергии вектор-контура в зависимости от размерности вектор-контура.



Как видно из рис.7, расхождение наблюдается только при малых размерностях пространства. Данное свойство можно использовать при расчете меры близости контуров.

Оценка выигрыша в производительности

Наличие фактора наследственности при эквализации контура позволяет обрабатывать контур в пространствах меньших размерностей, тем самым, экономя вычислительные ресурсы.

Предположим, что размерность контуров сокращается в a раз, т.е. $m = n/a$. Также учтем, что для эквализации предложенной в [1] необходимо $2n$ обращений к элементам контура. В связи с этим рассчитаем выигрыш W в трудоемкости алгоритма по обращению к элементам контура

$$W = n^2 - 2n - \left(\frac{n}{a}\right)^2$$

Отсюда

$$W = \frac{(a^2 - 1)n^2}{a^2} - 2n.$$

Для выигрыша в производительности необходимо $W > 0$.

Полученное неравенство с параметром a и неизвестным n , которое определяет эффективность вычислений, будет выглядеть следующим образом

$$\kappa(a)n - 2 > 0,$$

где $\kappa(a) = (a^2 - 1)/a^2$ - некоторый постоянный коэффициент, при заданном a .

Как видно, для определения эффективности эквализации контура к размерности $m = n/a$, необходимо решить линейное неравенство относительно n . С другой стороны, для заданного контура размерности n можно найти предельное значение A , которое является граничным, т.е. при $a > A$ эффективность преобразования теряется, а при $a < A$ вычисления являются эффективными.

Заключение

В данной статье исследуется фактор наследственности признака изображения «контур» при эквализации признака к различным размерностям. Выявляются зависимости значения меры близости контуров изображения, изменение энергии контура, влияние количества последовательных эквализаций контура при изменении размерности исходного эталонного контура. На основе полученных результатов исследования предлагается подход к обработке изображений и оценка эффективности применения предложенного подхода.

Литература

1. Фурман Я.А. Основы теории обработки контуров изображений: Учебное пособие. - Йошкар-Ола: МарГТУ, 1997. - 255с.

2. Жизняков А.Л. Теоретические основы обработки много-масштабных последовательностей цифровых изображений : монография / А.Л. Жизняков, С.С.Садыков; Владим. гос. ун-т.– Владимир: Изд-во Владим. гос. ун-та, 2008.- 121с.
3. Зуев В.В. Обработка изображений на основе базисов операторов. Математическое и программное обеспечение вычислительных систем: Межвуз. сб. науч. тр. / Под ред. А.Н. Пылькина – М.:Горячая линия – Телеком, 2009. 148с.

THE POSSIBILITIES OF THE FACTOR OF THE INHERITANCE OF ATTRIBUTES USING ON IMAGES SEQUENCES FOR THEIR ANALYSIS

V.V. Zuev, Murom Institute of the Vladimir State University

While images analyzing one of fundamental problem is high capacity of the input. Solving of this problem is processing the most informative part instead all image pixels. Such part is edges. For edges processing on the computer we must encode it. In the [1] paper some edges encoding types are reviewed including: a) Freeman code; b) two-dimensional code; 3) polygonal representation; 4) complex-valued code. The complex vector C^k space is considered the most universal for the edges comparing. The space is invariant to shift, scaling and rotation of edges. For edges comparing it is proposed to use the property of normed scalar multiplication (NSM); unit value for which means identity of comparing edges and zero one means full incongruence of them.

For that it is computed the cross-correlation function (CCF) of two analyzed edges and the maximum function value is taken as the closeness measure. The CCF computing operation is too much laborious and for space C^k amount of addressing to vector elements is k^2 . In this article the approach is proposed then allows reducing of addresses amount to edges elements as well as laboriousness of CCF-algorithm computing. Usually after the CC allocation it consists of elementary vectors (EV). Often that edge representation is redundant and we can describe the edge by smaller dimension CC without information losing or with insignificant losing. Equalization is the procedure of dimension alignment of two edges [1]. It can be used for reduction of the information.

In given article the opportunity of use of a heredity of a contour is analyzed at it equalization in spaces of various dimension on sequence of images.

In the [2] the theory that allows to prove the given similarity is described.

The inheritance presence at the edge equalization allows to process the edge in the spaces with fewer dimensions thereby economized computing resources.