

УДК 621.391.268

*Б.А. Алпатов, П.В. Бабаян, В.В. Стротов*

## **АНАЛИЗ ТОЧНОСТНЫХ ХАРАКТЕРИСТИК МЕТОДОВ СЛЕЖЕНИЯ ЗА ФОНОВЫМ ИЗОБРАЖЕНИЕМ ДЛЯ БОРТОВОЙ ВИДЕОИНФОРМАЦИОННОЙ СИСТЕМЫ**

*Рассмотрена проблема оценивания параметров сдвига и поворота фонового изображения в последовательности видеок кадров. Описано два алгоритма решения предлагаемой задачи. Приведены результаты сравнительных экспериментальных исследований этих алгоритмов. На основе исследований сформулированы рекомендации по использованию данных подходов в бортовых видеоинформационных системах.*

### **Введение**

Алгоритмы автоматического обнаружения объектов в последовательности изображений можно условно разделить по критерию характера производимой ими обработки на пространственные, временные и пространственно-временные. Среди них наилучшими параметрами обладают алгоритмы, производящие временной и пространственно-временной анализ последовательности видеоизображений [1]. Одним из требований, предъявляемых при проведении обработки изображений во времени, является знание пространственного соответствия между точками различных кадров видеопоследовательности. Это требование затрудняет применение временных и пространственно-временных алгоритмов обнаружения в бортовых системах видеонаблюдения и других системах обработки изображений с подвижным датчиком. К сожалению, использование таких средств, как измерение углов поворота или гиросtabilизация видеодатчиков с помощью специальных устройств, не всегда решает данную проблему. В связи с этим возникает задача оценки параметров геометрических преобразований в последовательности изображений на основе информации, заложенной в самой видеопоследовательности.

Существует большое количество методов решения поставленной задачи [2], основанных на корреляционном анализе, спектральном разложении, семантической интерпретации изображений. Выбор метода может зависеть от вида рассматриваемых преобразований, характера наблюдаемых изображений и наличия вычислительных ресурсов.

В бортовой видеоинформационной системе слежение за фоновым изображением должно выполняться в реальном масштабе времени, без вмешательства оператора и при отсутствии априорной информации о наблюдаемой сцене. В качестве допустимых видов преобразований рассматриваются сдвиг и поворот фонового изображения. Исходя из этих требований для исследований было выбрано два подхода к оценке геометрических преобразований изображения. В соответствии с первым подходом производится глобальное сравнение изображений, полученных с датчика в разные моменты времени [3, 4]. В основе алгоритма лежит известное утверждение о том, что поворот изображения ведёт к такому же повороту его амплитудного спектра, причём амплитудный спектр бесконечного изображения не зависит от его сдвига. Поскольку при сопоставлении изображений они рассматриваются как единые целые, этот подход далее будем называть одноэталонным. В соответствии со вторым, многоэталонным, подходом [2, 5] на изображении по определённому критерию выделяется ряд опорных участков небольших размеров. Отслеживая перемещения опорных областей, алгоритм одновременно оценивает геометрические эволюции, такие как сдвиг и поворот, происходящие в исходном изображении.

Экспериментальные исследования указанных алгоритмов были выполнены с использованием натуральных и искусственных видеопоследовательностей. Сравнение алгоритмов производилось по критерию среднеквадратической ошибки оценки параметров сдвига и поворота изображения. В результате исследований был

получен ответ на вопрос о целесообразности использования того или иного алгоритма в бортовой видеоинформационной системе.

**Постановка задачи**

Рассмотрим модель формирования изображения в видеоинформационной системе с подвижным датчиком. В качестве допустимых преобразований фонового изображения будем рассматривать только поворот и сдвиг, тогда можно ввести в рассмотрение непрерывное изображение  $q(x, y, t)$ ,  $(x, y) \in R^2$ , сформированное из изображений фона  $g(x, y, t)$  и объектов  $h(x, y, t)$ , где  $x, y$  - координаты точки,  $t$  - момент времени, в который производится наблюдение. В дальнейшем предполагается, что как объекты, так и фон являются полностью «непрозрачными» для диапазона частот, в котором принимается оптический сигнал. Атмосфера считается полностью прозрачной. При наблюдении за объектами, перемещающимися на сложном фоне, модель формирования изображения  $q(x, y, t)$  имеет вид:

$$q(x, y, t) = h(x, y, t)r(x, y, t) + g(x, y, t)(1 - r(x, y, t)), (x, y) \in R^2, \quad (1)$$

где функция  $r(x, y, t)$  определяет расположение объектов на изображении и задаётся правилом:

$$r(x, y, t) = \begin{cases} 1, & \text{если в точке } (x, y) \text{ в} \\ & \text{момент } t \text{ находится объект;} \\ 0 & \text{в противном случае.} \end{cases} \quad (2)$$

При наблюдении изображения в системе с подвижным видеодатчиком из  $q(x, y, t)$  выделяется прямоугольная область, которая затем дискретизируется в пространстве. Выделение прямоугольной области из изображения и её пространственная дискретизация происходят с частотой кадровой развёртки датчика изображений. Таким способом осуществляется дискретизация во времени. Предполагая, что интервалы дискретизации единичны и также учитывая аддитивный шум, вносимый датчиком изображения, получаем следующую модель формирования дискретного наблюдаемого изображения:

$$l(i, j, n) = q(\lambda(i, j, n), \mu(i, j, n), n) + \xi(i, j, n), (i, j) \in X, n = \overline{1, N}. \quad (3)$$

Здесь и далее  $i, j$  - координаты пикселя,  $(i, j) \in X$ ,

$$X = \left\{ (i, j) : i = \overline{-(N_x - 1)/2, (N_x - 1)/2}, j = \overline{-(N_y - 1)/2, (N_y - 1)/2} \right\}$$

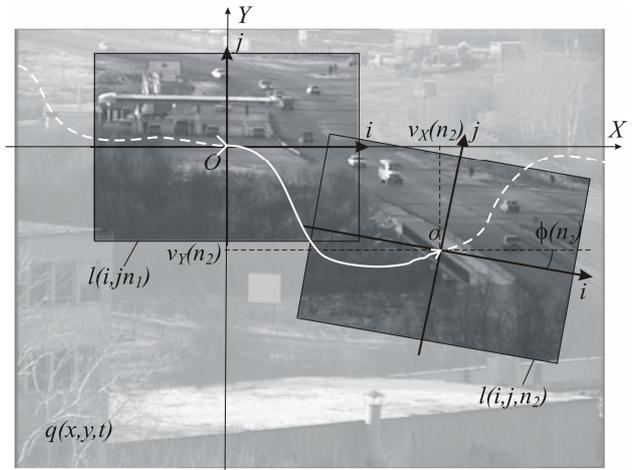
где  $X$  - множество точек, на котором задано изображение;  $N_x, N_y$  - размеры изображе-

ния (для упрощения дальнейшего изложения полагаются нечётными);  $\xi(x, y, n)$  - шум с нулевым средним;  $l(i, j, n)$  - наблюдаемое изображение,

$$\begin{cases} \lambda(i, j, n) = i \cdot \cos \varphi(n) + j \cdot \sin \varphi(n) - v_x(n); \\ \mu(i, j, n) = -i \cdot \sin \varphi(n) + j \cdot \cos \varphi(n) - v_y(n) \end{cases}$$

функции, описывающие движение прямоугольной области по изображению  $q(x, y, t)$ . Функция  $\varphi(n)$  задаёт ориентацию данного кадра относительно оси  $OX$ , а функции  $v_x(n), v_y(n)$  определяют сдвиг центра прямоугольной области.

Суть сказанного ясна из рассмотрения рисунка 1. Здесь центр поля зрения датчика двигается вдоль некоторой траектории (на рисунке - белая кривая), и одновременно изменяется ориентация датчика, то есть поле зрения поворачивается. При этом происходят соответствующие изменения в наблюдаемых датчиком изображениях. Некоторые части изображения уходят из поля зрения, некоторые, наоборот, входят. Остальные области остаются в поле зрения, но подвергаются преобразованиям сдвига и поворота.



**Рис. 1. Формирование изображения при подвижном датчике**

Одной из проблем, возникающих при решении задачи, является наличие подвижных объектов, движение которых независимо от движения фона.

Далее будет рассматриваться задача оценки сдвига и поворота фоновых участков двух изображений  $l_1(i, j)$  и  $l_2(i, j)$ , полученных датчиком в разные моменты времени на кадрах с номерами  $n_1$  и  $n_2$ :  $l_1(i, j) = l(i, j, n_1)$ ,  $l_2(i, j) = l(i, j, n_2)$ . Считается, что фон не изменяется от кадра к кадру, то есть в выражении (1)

$$g(i, j, n) = g(i, j, n - 1) = g(i, j). \quad (4)$$

Выберем систему координат для изображения  $q(i, j, n)$  таким образом, чтобы выполнялись условия:

$$\varphi(n_1) = 0; \quad v_x(n_1) = 0; \quad v_y(n_1) = 0.$$

Тогда модель формирования изображений  $l_1(i, j)$  и  $l_2(i, j)$  выглядит следующим образом:

$$\begin{aligned} l_1(i, j) &= g(i, j)(1 - r_1(i, j)) + h_1(i, j)r_1(i, j) + \xi_1(i, j), \\ l_2(i, j) &= g(\lambda(i, j, n), \mu(i, j, n))(1 - r_2(\lambda(i, j, n), \mu(i, j, n))) + \\ &+ h_2(\lambda(i, j, n), \mu(i, j, n))r_2(\lambda(i, j, n), \mu(i, j, n)) + \xi_2(i, j), \end{aligned} \quad (5)$$

$$(i, j) \in X,$$

где  $r_1(i, j) = r(i, j, n_1)$ ,  $r_2(i, j) = r(i, j, n_2)$  – матрицы, задающие расположение объектов для изображений  $l_1(i, j)$  и  $l_2(i, j)$  соответственно;  $h_1(i, j) = h(i, j, n_1)$  и  $h_2(i, j) = h(i, j, n_2)$  – яркостные изображения объектов, присутствующих на  $l_1(i, j)$  и  $l_2(i, j)$ ;  $\xi_1(i, j) = \xi(i, j, n_1)$ ,  $\xi_2(i, j) = \xi(i, j, n_2)$  – соответствующие реализации шума.

Задача состоит в том, чтобы оценить  $v_x(n_2)$ ,  $v_y(n_2)$  и  $\varphi(n_2)$  на основе известных изображений  $l_1(i, j)$ ,  $l_2(i, j)$ ,  $(i, j) \in X$ , модель формирования которых задана уравнениями (5).

### Одноэталонный алгоритм слежения за фоном

Одноэталонный подход к нахождению параметров геометрических преобразований фона, в отличие от многоэталонного, предполагает рассмотрение изображения в целом, без выделения из него характерных областей. Для этих целей мог бы использоваться обычный корреляционный метод, в соответствии с которым производится перебор всех допустимых преобразований изображения и поиск максимального значения некоторой функции сходства [2]. Недостатком корреляционного алгоритма является большая вычислительная сложность. Этот недостаток становится ещё более ощутимым, во-первых, из-за необходимости анализа изображений больших размеров, и, во-вторых, в связи с тем, что пространство допустимых преобразований является трёхмерным:  $(v_x, v_y, \varphi)$ . Поскольку разрабатываемый алгоритм слежения должен работать в режиме реального времени, обычные корреляционные алгоритмы здесь неприемлемы. Одним из хороших методов снижения вычислительных затрат при реализации корреляционных алгоритмов является использование быстрых спектральных преобразований. Одноэталонный алгоритм, описанный в этой работе, использует быстрое преобразование Фурье для оценки параметров поворота и сдвига фона.

В основе одноэталонного алгоритма слежения за фоном лежат теоремы смещения и вращения для спектров функций двух переменных. Пусть  $f_1(x, y)$  и  $f_2(x, y)$  – функции, определённые на плоскости  $R^2$ , которые можно рассматривать как бесконечные изображения, а  $F_1(\omega_x, \omega_y)$  и  $F_2(\omega_x, \omega_y)$  – преобразования Фурье от этих функций. Теорема смещения утверждает, что если  $f_1(x, y)$  и  $f_2(x, y)$  связаны преобразованием сдвига

$$f_2(x, y) = f_1(x - v_x, y - v_y), \quad (6)$$

где  $(v_x, v_y)$  – вектор сдвига, то их спектры связаны соотношением

$$F_2(\omega_x, \omega_y) = e^{-2\pi j(\omega_x v_x + \omega_y v_y)} F_1(\omega_x, \omega_y), \quad (7)$$

где  $j = \sqrt{-1}$ .

Из теоремы смещения легко получить алгоритм нахождения вектора  $(v_x, v_y)$ . Из (7) получим фазовую составляющую взаимного энергетического спектра функций  $f_1(x, y)$  и  $f_2(x, y)$ :

$$e^{2\pi j(\omega_x v_x + \omega_y v_y)} = \frac{F_1(\omega_x, \omega_y) F_2^*(\omega_x, \omega_y)}{|F_1(\omega_x, \omega_y) F_2(\omega_x, \omega_y)|}.$$

Выполнив обратное преобразование Фурье от последнего выражения, получим двумерный дельта-импульс

$$\delta(x - v_x, y - v_y) = \Phi^{-1} \left\{ e^{2\pi j(\omega_x v_x + \omega_y v_y)} \right\}.$$

Положение максимума дельта-импульса соответствует вектору сдвига  $(v_x, v_y)$ . При обработке реальных изображений, искаженных шумом, вместо дельта-импульса мы получим функцию, которая близка к нулю везде, кроме окрестности точки  $(v_x, v_y)$ . Метод оценки параметров сдвига изображений, основанный на поиске точки максимума этой функции, называется алгоритмом фазовой корреляции.

Теорема вращения утверждает, что если функции  $f_1(x, y)$  и  $f_2(x, y)$  связаны преобразованием поворота на угол  $\varphi$ , то есть

$$f_2(x, y) = f_1(x \cos \varphi + y \sin \varphi, -x \sin \varphi + y \cos \varphi), \quad (8)$$

то спектры этих функций связаны соотношением

$$\begin{aligned} F_2(\omega_x, \omega_y) &= F_1(\omega_x \cos \varphi + \omega_y \sin \varphi, \\ &- \omega_x \sin \varphi + \omega_y \cos \varphi). \end{aligned} \quad (9)$$

Таким образом, поворот функции на угол  $\varphi$  соответствует точно такому же повороту её спектра.

Из (6)–(9) следует, что если функция  $f_2(x, y)$  получена путём сдвига и поворота функции  $f_1(x, y)$ , то есть

$$f_2(x, y) = f_1(x \cos \varphi + y \sin \varphi - v_x, -x \sin \varphi + y \cos \varphi - v_y), \quad (10)$$

то

$$F_2(\omega_x, \omega_y) = e^{-2\pi j(\omega_x v_x + \omega_y v_y)} \cdot F_1(\omega_x \cos \varphi + \omega_y \sin \varphi, -\omega_x \sin \varphi + \omega_y \cos \varphi). \quad (11)$$

Пусть  $M_1(\omega_x, \omega_y)$  и  $M_2(\omega_x, \omega_y)$  – амплитудные спектры функций  $f_1(x, y)$  и  $f_2(x, y)$ , то есть  $M_1(\omega_x, \omega_y) = |F_1(\omega_x, \omega_y)|$ ,

$$M_2(\omega_x, \omega_y) = |F_2(\omega_x, \omega_y)|. \quad \text{Из (11) следует}$$

$$M_2(\omega_x, \omega_y) = M_1(\omega_x \cos \varphi + \omega_y \sin \varphi, -\omega_x \sin \varphi + \omega_y \cos \varphi). \quad (12)$$

Таким образом, при сдвиге и повороте функции происходит такой же поворот её амплитудного спектра, а сдвиг не оказывает влияния на амплитудный спектр функции. Переходя к амплитудным спектрам, мы освобождаемся от влияния сдвигов и можем отслеживать поворот отдельно от других видов преобразований.

Для этого удобно в выражении (12) перейти в полярную систему координат:  $\omega_x = \rho \cos \theta$ ,  $\omega_y = \rho \sin \theta$ . В полярных координатах амплитудные спектры связаны преобразованием сдвига вдоль оси  $\theta$ :

$$M_2(\rho, \theta) = M_1(\rho, \theta - \varphi). \quad (13)$$

Угол  $\varphi$  легко находится методом фазовой корреляции, изложенным выше. Зная  $\varphi$ , можно развернуть функцию  $f_1(x, y)$  на угол  $\varphi$ . Тогда мы получим две функции, связанные преобразованием сдвига. Применяя вновь метод фазовой корреляции, получаем оценки параметров сдвига  $(v_x, v_y)$ .

Отметим, что существует аналогичный подход, позволяющий находить, наряду со сдвигом и поворотом, ещё и параметр изменения масштаба [3, 4].

Приведённые выше соотношения справедливы для изображений, рассматриваемых на непрерывной плоскости  $R^2$ , в отсутствие шумов и движущихся объектов. Но, тем не менее, сделанные выводы в первом приближении имеют место и в натуральных условиях наблюдения.

Рассмотрим алгоритм оценки параметров сдвига и поворота изображений  $l_1(i, j)$ ,  $l_2(i, j)$ ,  $(i, j) \in X$  применительно к модели (5). Для уп-

рощения изложения будем предполагать, что рассматриваемое изображение является квадратным размером  $N \times N$  ( $N_x = N$ ,  $N_y = N$ ).

Вместо непрерывного преобразования Фурье здесь необходимо использовать дискретное преобразование Фурье (ДПФ).

При этом возникает ряд проблем. Во-первых, амплитудный спектр дискретного изображения оказывается сильно искаженным вследствие хорошо известного явления Гиббса. Рассмотрим два изображения и логарифмы их амплитудных спектров (рис. 2). Логарифмирование применено здесь в иллюстративных целях, чтобы улучшить различимость высокочастотных составляющих, имеющих малую амплитуду.

Хорошо видно наличие светлого креста на изображениях логарифма спектра. Ориентация этого креста не зависит от поворота изображения, а амплитуда сравнима с амплитудой других составляющих спектра. Возникновение креста связано с тем, что при выполнении ДПФ изображение рассматривается как периодическое. Скачки яркости, присутствующие на границах периодов изображения, вносят в спектр дополнительные частотные составляющие. Этот эффект называют явлением Гиббса.

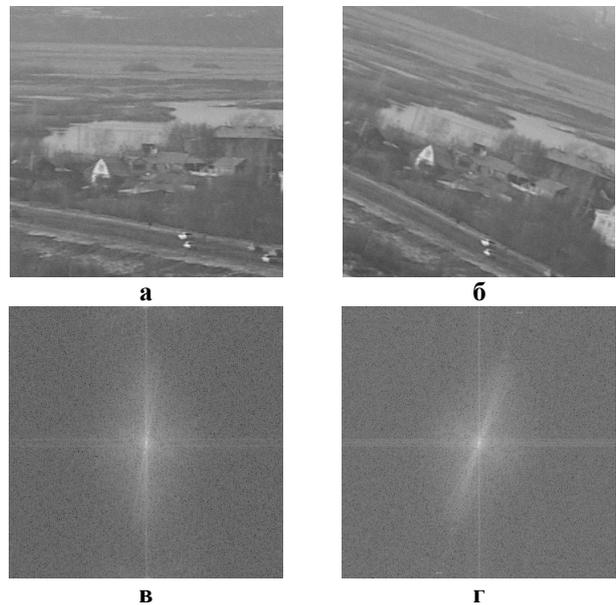


Рис. 2. Изображения, отличающиеся поворотом (а, б), и логарифмы их амплитудных спектров (в, г)

Поскольку скачки яркости расположены на границах прямоугольной области  $X$ , в спектр включаются компоненты, связанные с конфигурацией множества  $X$ . Несмотря на то, что спектр, тем не менее, содержит составляющую, поворачивающуюся вместе с изображением, присутствие неподвижного креста мешает найти точку наилучшего совмещения спектров. Чтобы

избавиться от креста, можно использовать стандартные приёмы уменьшения явления Гиббса: умножение изображения на окна Хэмминга, Ханна, Чебышева, Кайзера [6]. Здесь предлагается более простой подход, при котором осуществляется заполнение средним значением яркости участков изображения, не принадлежащих кругу с центром в начале координат и радиусом  $(N-1)/2$ :

$$l^w(i, j) = \begin{cases} E_{(i,j) \in X} \{l(i, j)\}, & \text{если } \sqrt{i^2 + j^2} > (N-1)/2; \\ l(i, j) & \text{в противном случае,} \end{cases} \quad (14)$$

где  $E$  – оператор усреднения,  $l^w(i, j)$  – преобразованное изображение. Использование формулы (15), по существу, не устраняет явление Гиббса, поскольку скачки яркости всё ещё присутствуют в изображении. Однако скачки яркости находятся на границе круглой области, имеющей центральную симметрию, что значительно сокращает отрицательное влияние эффекта Гиббса. Рассмотрим изображение (рис. 2,а), умноженное на окно Кайзера (рис. 3,а), а также подвергнутое преобразованию (14) (рис. 3,б).

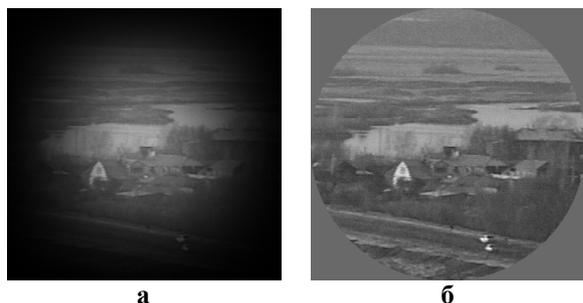


Рис. 3. Изображение, умноженное на окно Кайзера (а), преобразованное по формуле (15) (б)

Как видно из рисунка, использование окна Кайзера, как и других видов окон, упомянутых выше, ведёт к потере некоторой части яркостной информации, находящейся на краю изображения.

Вторая проблема, возникающая при анализе реальных изображений, связана с недостаточным уровнем высокочастотных составляющих изображения. Из-за этого точность определения угла поворота  $\varphi$  оказывается низкой. Чтобы решить эту проблему, в [4] был использован фильтр высоких частот с передаточной функцией:

$$H(u, v) = (1 - Q(u, v)) \cdot (2 - Q(u, v)), \quad (15)$$

где  $Q(u, v) = \cos\left(\pi \frac{u}{(N-1)/2}\right) \cos\left(\pi \frac{v}{(N-1)/2}\right);$   
 $(u, v) \in X$ .

В [3] для тех же целей применялось логарифмирование амплитудного спектра. Поскольку амплитуда высокочастотных составляющих обычно ниже, чем у низкочастотных, путём логарифмирования фактически производится высокочастотная фильтрация изображения.

Третья проблема связана с использованием метода фазовой корреляции для определения вектора  $(v_x, v_y)$  и угла  $\varphi$ . Метод фазовой корреляции (8, 9) хорошо работает лишь при малом уровне искажений и шумов [7]. В присутствии шума и движущихся объектов лучше ведут себя корреляционный и разностный алгоритмы. В данной работе использовался корреляционный алгоритм, который отличается от фазового корреляционного отсутствием деления на модуль произведения спектров функций в выражении (8). В результате обратного преобразования Фурье мы получим взаимную корреляционную функцию, максимум которой укажет нам значенные параметров смещения.

### Многоэталонный алгоритм слежения за фоном

В основе многоэталонного алгоритма лежит принцип пространственно-временной обработки изображений, включающий выбор на изображении опорных участков и определение их положения от кадра к кадру. На основании оценки положения этих участков производится оценка параметров поворота и смещения изображения. В качестве эталона на изображении выбирается неподвижный участок фона.

В начальный момент времени выбираются  $M_{эм}$  участков исходного изображения (эталон), по положению которых определяются смещение и поворот фона. Обозначим  $m$ -й эталон

$w_m(i, j, n)$  ( $i \in \overline{1, a_m}, j \in \overline{1, b_m}, m = \overline{1, M_{эм}}, n$  – номер кадра). Количество и размер опорных участков могут варьироваться в зависимости от характеристик видеопоследовательности. Идеальным опорным участком будет изображение неподвижного объекта, при использовании которого корреляционный алгоритм измерения координат будет давать наименьшие ошибки (например, угол здания или здание с окнами). Критериями выбора эталона будут его положение (не слишком близко к краям изображения) и хорошая отличимость от окружающих его фоновых образований.

Один из подходов к выбору опорных участков состоит в том, что эталон должен иметь максимальную дисперсию среди всех участков в области выбора, т.е.

$$(i_m^h, j_m^h) = \arg \max_{i,j \in Z_m} \left( \frac{1}{a_m \cdot b_m} \sum_{\alpha=i-\frac{a_m}{2}}^{i+\frac{a_m}{2}-1} \sum_{\beta=j-\frac{b_m}{2}}^{j+\frac{b_m}{2}-1} (l_1(\alpha, \beta) - \bar{l}_1^{i,j})^2 \right), \quad (16)$$

где  $i_m^h, j_m^h$  - начальные координаты  $m$ -го эталона ( $m = \overline{1, M_{эм}}$ ),  $Z_m$  - зона выбора  $m$ -го эталона,

$$\bar{l}_1^{i,j} = \frac{1}{a_m \cdot b_m} \sum_{\alpha=i-\frac{a_m}{2}}^{i+\frac{a_m}{2}-1} \sum_{\beta=j-\frac{b_m}{2}}^{j+\frac{b_m}{2}-1} l_1(\alpha, \beta) - \text{среднее}$$

значение яркости по участку изображения.

Возможны также другие подходы к выбору эталонов, например по максимуму отличия от участков в ближайшей окрестности [5].

Зона выбора эталона представляет собой сектор, образованный углом величиной  $360^\circ / M_{эм}$ . Радиус сектора равен половине минимального из линейных размеров наблюдаемого изображения, причём центр окружности совмещён с центром изображения.

Примеры выбранных алгоритмом опорных участков представлены на рис. 4 ( $M_{эм} = 4$ ).

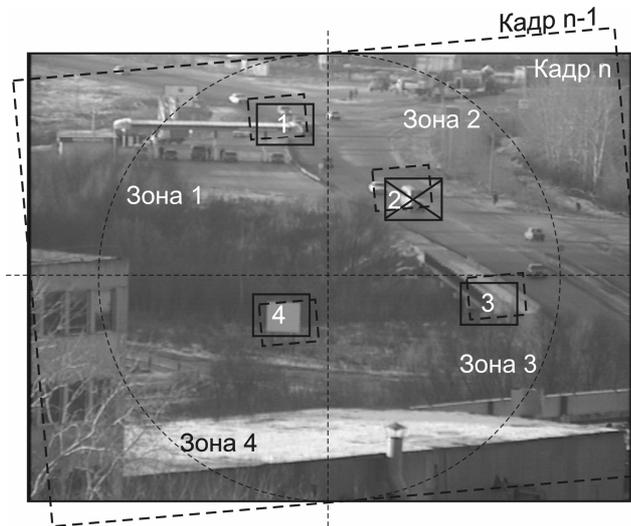


Рис. 4. Пример работы многоэталонного алгоритма для  $M_{эм} = 4$

Оценка положения эталонов в текущем кадре  $n$  производится с использованием корреляционно – экстремальных методов [9]. Координаты  $m$ -го эталона  $w_m$  могут быть найдены как:

$$\begin{aligned} (\hat{i}_m(n), \hat{j}_m(n)) &= \arg \min_{i,j \in W_m(n)} (F_{крум}(i, j)), \text{ где} \\ F_{крум}(i, j) &= \sum_{\alpha=\beta=1}^{a_m, b_m} \left| l_1\left(i + \alpha - \frac{a_m}{2}, j + \beta - \frac{b_m}{2}, n\right) - \right. \\ &\quad \left. - w_m(\alpha, \beta, n) \right|. \end{aligned} \quad (17)$$

Здесь  $W_m(n)$  - зона поиска эталона  $w_m$  в  $n$ -м кадре. Она представляет собой прямоугольную область, по которой движется центр эталона. Положение данной области определяется исходя из значений  $i_m, j_m$  и оценок функций  $v_x(n), v_y(n), \varphi(n)$  в предыдущем кадре  $\hat{v}_x(n-1), \hat{v}_y(n-1), \hat{\varphi}(n-1)$  (она располагается в точке прогнозируемого расположения объекта), её размеры также пропорциональны оценкам функций смещения и поворота.

Оценивание координат производится с точностью до целого пикселя, однако возможно субпиксельное уточнение найденных значений [10]. Уточняющие субпиксельные поправки для оценок  $\hat{i}_m(n), \hat{j}_m(n)$  могут быть получены параболическим интерполированием разностной критериальной функции  $F_{крум}$  в ближайшей окрестности точки минимума.

После определения и уточнения координат всех опорных участков на основе модели преобразования составляется система уравнений, связывающая положение эталонов в текущем и предшествующем кадрах:

$$\begin{bmatrix} \hat{i}_1(n+1) \\ \hat{j}_1(n+1) \\ \dots \\ \hat{i}_{M_{эм}}(n+1) \\ \hat{j}_{M_{эм}}(n+1) \end{bmatrix} = \begin{bmatrix} \hat{i}_1(n) & \hat{j}_1(n) & 1 & 0 \\ \hat{j}_1(n) & -\hat{i}_1(n) & 0 & 1 \\ \dots & \dots & \dots & \dots \\ \hat{i}_{M_{эм}}(n) & \hat{j}_{M_{эм}}(n) & 1 & 0 \\ \hat{j}_{M_{эм}}(n) & -\hat{i}_{M_{эм}}(n) & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos(\Delta\varphi(n)) \\ \sin(\Delta\varphi(n)) \\ \Delta v_x(n) \\ \Delta v_y(n) \end{bmatrix}, \quad (18)$$

где  $\Delta\varphi(n) = \varphi(n) - \varphi(n-1)$  - угол, на который повернулось наблюдаемое изображение вокруг своего центра за время формирования  $n$ -го кадра,  $\Delta v_x(n) = v_x(n) - v_x(n-1), \Delta v_y(n) = v_y(n) - v_y(n-1)$  - смещение наблюдаемого изображения за время формирования  $n$ -го кадра.

Ввиду того что количество опорных участков больше двух, система является переопределённой ( $2M_{эм} \times 4$ ), то есть в общем случае не имеющей точного решения. В этом случае возможно нахождение такого решения, которое обеспечит минимум среднеквадратической ошибки. Такое решение может быть получено

при применении метода псевдообращения матрицы [8]:

$$\begin{bmatrix} \cos(\hat{\Delta\varphi}(n)) \\ \sin(\hat{\Delta\varphi}(n)) \\ \hat{\Delta v}_x(n) \\ \hat{\Delta v}_y(n) \end{bmatrix} = \begin{bmatrix} \hat{i}_1(n) & \hat{j}_1(n) & 1 & 0 \\ \hat{j}_1(n) & -\hat{i}_1(n) & 0 & 1 \\ \dots & \dots & \dots & \dots \\ \hat{i}_{M_{эм}}(n) & \hat{j}_{M_{эм}}(n) & 1 & 0 \\ \hat{j}_{M_{эм}}(n) & -\hat{i}_{M_{эм}}(n) & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \hat{i}_1(n+1) \\ \hat{j}_1(n+1) \\ \dots \\ \hat{i}_{M_{эм}}(n+1) \\ \hat{j}_{M_{эм}}(n+1) \end{bmatrix} \quad (19)$$

где  $A^-$  - матрица, псевдообратная к прямоугольной матрице  $A$ .

В общем случае некоторые выбранные по критерию (16) опорные участки могут оказаться подвижными (как, например, эталон 2 на рис. 4), поэтому полученные оценки параметров преобразований могут существенно отличаться от истинного значения, а значит, следующим этапом работы алгоритма является уточнение оценок. Для этого вычисляется столбец невязок  $[d_m(n)]$  для данного решения, а затем составляется новая система, отличающаяся от системы (17) тем, что в ней отсутствуют строки, для которых  $d_m(n) < d_{порога}(m = \overline{1, 2 \cdot M_{эм}})$ . Это делается для того, чтобы исключить влияние на оценку параметров грубых ошибок определения координат, а также убрать из рассмотрения движущиеся опорные участки. Величина  $d_{порога}$  выбирается равной 0,5 пикс., однако, в зависимости от ситуации, она может быть увеличена. Результатом решения данной системы являются уточнённые

$$\hat{\Delta v}_x, \hat{\Delta v}_y, \hat{\Delta\varphi} = \arctan \frac{\sin(\hat{\Delta\varphi}(n))}{\cos(\hat{\Delta\varphi}(n))}.$$

### Сравнительное исследование алгоритмов

Для сравнительного исследования алгоритмов было создано 24 видеопоследовательности размером  $339 \times 339$  пикселей. Половина из них не содержала движущихся объектов и была получена из одного исходного изображения размером  $768 \times 576$ . Эти 12 сюжетов назовём статическими. Другая половина сюжетов содержала движущиеся объекты, занимающие около 5% изображения, и получена из видеопоследовательности с размером кадра  $768 \times 576$ , снятой при неподвижном датчике изображения и содержащей движущиеся объекты. Назовём эти сюжеты динамическими. Процесс получения каждой видеопоследовательности можно представить как процесс движения по исходному изображению прямоугольного окна размером

$339 \times 339$  пикселей. Угол наклона прямоугольного окна изменялся от кадра к кадру с приращением в +1 или -1 градус. Максимальное значение модуля этого угла составляло 0, 1, 5, 10, 45, 90 градусов. По достижении модулем угла максимального значения знак приращения угла менялся на противоположный. Центр прямоугольного окна двигался по окружности, центр которой находился в центре изображения, а радиус окружности в разных сюжетах принимал значения 0 или 10 пикселей.

Полученные видеосюжеты обрабатывались одноэталонным (ОЭА) и многоэталонным (МЭА) алгоритмами, затем вычислялись СКО для каждой видеопоследовательности по углу поворота  $\varphi$  и одной из координат  $x$ . Результаты для всех сюжетов приведены на рис. 5.

Из рассмотрения рис. 5 вытекает, что одноэталонный алгоритм даже при малых углах поворота позволяет оценить параметры сдвигов и поворотов гораздо точнее, причём с ростом максимального угла поворота разница в точности между двумя алгоритмами возрастает.

Это позволяет говорить о том, что из двух предложенных подходов в видеоинформационной системе сопровождения объектов целесообразнее использовать одноэталонный. Многоэталонный алгоритм в том виде, в котором он существует на данный момент, не позволяет добиться приемлемой точности, что делает затруднительным использование его для стабилизации в алгоритмах выделения движения объектов. Однако его можно использовать для стабилизации изображения в системах улучшения видеона.

### Заключение

В статье рассмотрены два подхода к решению задачи слежения за смещающимся и поворачивающимся фоновым изображением в присутствии мешающих движущихся объектов. Первый метод, одноэталонный, производит анализ изображения в целом. В соответствии со вторым, многоэталонным, на изображении ищутся опорные участки. Отслеживая эти участки, алгоритм вычисляет параметры движения фонового изображения. Экспериментальные исследования, выполненные на большом количестве сюжетов, показали, что одноэталонный алгоритм имеет лучшие характеристики и может быть успешно использован в составе бортовой компьютерной системы обнаружения и слежения за движущимися объектами. Кроме того, обе рассмотренных методики могут быть успешно применены в устройствах визуальной электронной стабилизации изображений.

**Работа выполнена при поддержке гранта РФФИ № 06-01-08021.**

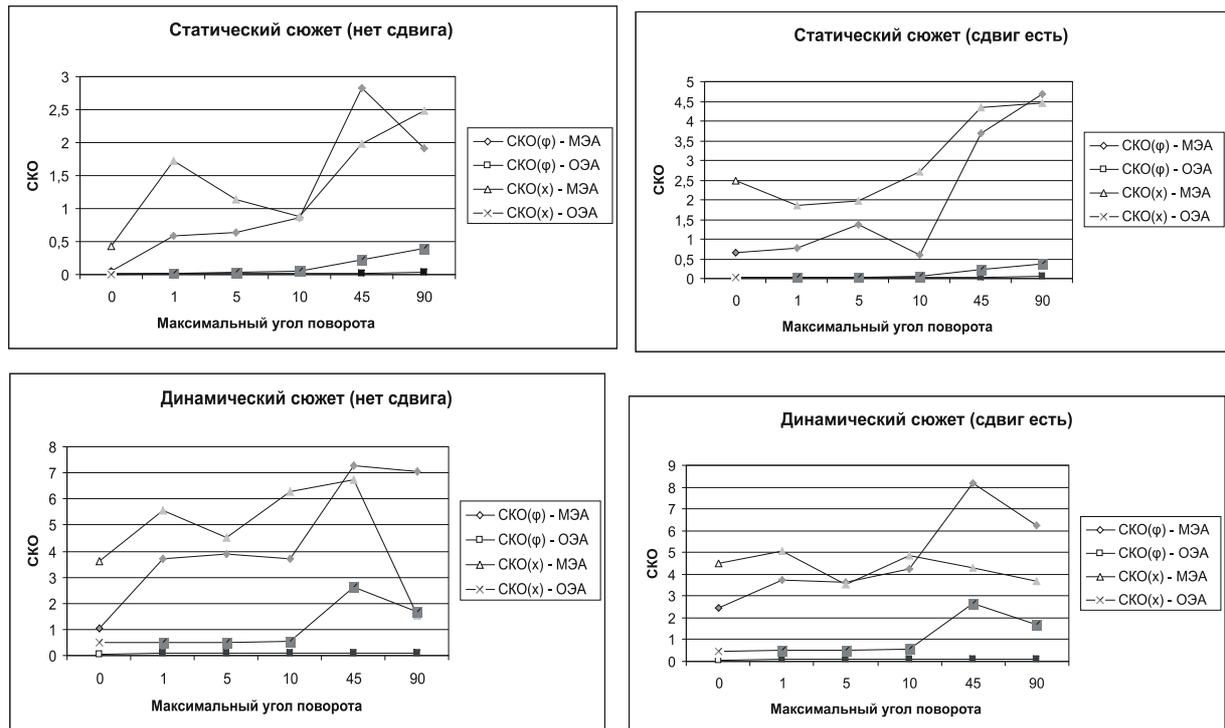


Рис. 5. Результаты исследований алгоритмов

### Библиографический список

1. Алпатов Б.А., Бабаян П.В. Выделение движущихся объектов в условиях геометрических искажений изображения // Цифровая обработка сигналов. – 2004. – №4. – С. 9-14.
2. L. G. Brown. A survey of image registration techniques // ACM Computing Surveys. Vol. 24. No. 4. – Dec. 1992. Pp. 325-376.
3. Murali Rao. Target recognition using cepstrum and inverse filtering // Proc. of SPIE. Vol. 2484. Signal Processing, Sensor Fusion and Target Recognition IV. – 1995. Pp. 224-235.
4. B. S. Reddy and B. N. Chatterji. An FFT-based Technique for Translation, Rotation and Scale-Invariant Image Registration // IEEE Transactions on Image Processing. Vol. 5, No. 8. – August, 1996. Pp. 1266-1271.
5. Бачило С.А., Дзягун Д.Ю., Итенберг И.И., Овчаров А.Б., Сивцов С.А. Средства электронной стабилизации телевизионных изображений для подвижных

- систем наблюдения. // Тез. докл. 4-й конф. DSPA. – М.: МЦНТИ 2002. – С. 252-254.
6. Adam Wilmer. Fourier-Mellin based Image Registration (with GUI). <http://www.mathworks.com/matlabcentral/fileexchange> (20.10.2004). Программное обеспечение для Matlab.
7. Баклицкий В.К., Бочкарёв А.М. Методы фильтрации сигналов в корреляционно-экстремальных системах навигации. – М.: Радио и связь, 1986. – 216 с.
8. Форсайт Дж., Малькольм М., Моулер К. Машинные методы математических вычислений. – М.: Мир, 1980. – 278 с.
9. Алпатов Б. А. Оптимальное оценивание параметров движущегося объекта в последовательности изображений // Автометрия. – 1994. – №2. – С. 32-37.
10. Деренченко Л.П., Троицкий В.И. Повышение точности корреляционно-экстремальных систем методом интерполяции целевой функции // Вопросы радиоэлектроники: Сб. науч. тр. – М., 1992. – с. 16-18.