

**В. С. Киричук, А. К. Шакенов***(Новосибирск)***АЛГОРИТМЫ ОБНАРУЖЕНИЯ ТОЧЕЧНЫХ ОБЪЕКТОВ  
ПО СТЕРЕОИЗОБРАЖЕНИЯМ**

Рассмотрены различные подходы к повышению качества обнаружения точечных динамических объектов за счет использования стереовидения. Реализованы и исследованы триангуляционный алгоритм с предварительным пороговым отсечением и алгоритм, формирующий множество отметок по всей совокупности стереоданных. При разных соотношениях полезного сигнала и шума для предложенных алгоритмов экспериментальным путем получена зависимость вероятности обнаружения от уровня ложных тревог. Приведена теоретическая зависимость амплитуды объекта от уровня ложной тревоги при заданной вероятности обнаружения. Исследована эффективность использования априорной информации о высотном ходе объекта в процессе его обнаружения.

**Введение.** Задача надежного обнаружения слабоконтрастных малоразмерных динамических объектов при малом соотношении полезного сигнала и шума крайне сложна. Естественным решением улучшения качества обнаружения является использование дополнительной информации, а именно синхронное наблюдение объекта с различных точек пространства. Подходы к применению стереовидения в задачах обнаружения и сопровождения малоразмерных движущихся объектов начали разрабатываться относительно недавно. Так, в работе [1] рассмотрен алгоритм для обнаружения объектов по стереопоследовательности, полученной стационарными камерами. Целью данной работы является подробное исследование различных алгоритмов (алгоритма монообработки, триангуляционного алгоритма, алгоритма стереообработки совместно по двум кадрам) с использованием двух изображений, полученных движущимися камерами, для улучшения качества обнаружения объекта.

**Постановка задачи.** Пусть две одинаковые синхронизированные камеры на протяжении некоторого периода времени наблюдают один и тот же движущийся объект. Предполагается, что геометрия съемки известна и стационарная часть фона оценена и устранена соответствующими процедурами межкадровой обработки (МКО) независимо по каждой из монопоследовательностей. Таким образом, входные последовательности содержат изображение объекта и случайный шум, обусловленный помехами регистрации и ошибками в компенсации стационарной части фона. Требуется обнаружить объект по данной стереопоследовательности.

**Математическая модель.** Пусть изображения  $D_m^k(x_i, y_j)$  допускают следующее описание:

$$D_m^k(x_i, y_j) = \psi_m^k(x_i, y_j) + A_m \varphi(X_m^k - x_i, Y_m^k - y_j), \quad (1)$$

где  $i = 1, \dots, N_x$ ;  $j = 1, \dots, N_y$ ;  $N_x, N_y$  – размеры изображений;  $m = 1, 2$  – номер последовательности;  $k = 1, \dots, T$  – номер кадра в последовательности;  $T$  – количество кадров;  $\psi_m^k(x, y)$  – случайная величина с нормальным распределением  $N(\mu_m^k, (\sigma_m^k)^2)$ ;  $A_m$  – амплитуда объекта;  $\varphi$  – форма объекта;  $(X_m^k, Y_m^k)$  – координаты объекта. Считая, что  $\mu_m^k \approx 0$ ,  $\sigma_m^k \approx \sigma$ ,  $A_1 = A_2$ , и проведя нормировку, получаем

$$D_m^k(x_i, y_j) = \psi(x_i, y_j) + A \varphi(X_m^k - x_i, Y_m^k - y_j), \quad (2)$$

где  $\psi(x_i, y_j)$  – случайная величина с нормальным распределением  $N(0, 1)$ ;  $A = A_1 / \sigma$ .

**Алгоритм монообработки.** При монообработке в каждой точке  $(x_i, y_j)$  изображения  $D_m^k(x_i, y_j)$  проверяется гипотеза о наличии объекта. Яркость в точке  $(X_m^k, Y_m^k)$ , соответствующей реальному положению объекта, имеет распределение  $N(A, 1)$ , а во всех остальных точках –  $N(0, 1)$ . Оптимальным критерием проверки двух простых гипотез является критерий отношения правдоподобия [2], который в данном случае может быть представлен в виде

$$D_m^k(x_i, y_j) > C_M, \quad (3)$$

где значение порога  $C_M$  определяется вероятностью ложной тревоги. Пусть  $\alpha_M$  – заданная вероятность ложной тревоги. Тогда по определению для  $(x_i, y_j) \neq (X_m^k, Y_m^k)$  имеем  $P(D_m^k(x_i, y_j) > C_M) = \alpha_M$ , откуда

$$C_M = \Phi_{(0, 1)}^{-1}(1 - \alpha_M), \quad (4)$$

где  $\Phi_{(\mu, \sigma^2)}(t)$  – функция нормального распределения  $N(\mu, \sigma^2)$ . Пусть  $\beta_M$  – вероятность пропуска объекта. Тогда очевидно, что

$$\beta_M = \Phi_{(A, 1)}(C_M). \quad (5)$$

Обнаружение по монопоследовательности не дает информации о трехмерных координатах объекта.

**Триангуляционный алгоритм обнаружения.** На первом этапе алгоритма формируются отметки (точки, в которых выполняется соотношение (3)) в каждом кадре стереопары. На втором этапе обработки полученное множество отметок уменьшается за счет применения стереоэффекта. Для оставшихся отметок вычисляются трехмерные координаты. Рассмотрим стереопару изображений  $D_1^k(x_i, y_j)$  и  $D_2^k(x_i, y_j)$ , полученных разными камерами в один момент времени. Введем обозначения  $\mathbf{p} = (x, y)$  для точек изображения  $D_1^k(x, y)$  и  $\mathbf{q} = (x, y)$  для точек изображения  $D_2^k(x, y)$ . Пронумеровав некото-

рым образом узлы дискретной решетки, введем обозначения  $\mathbf{p}_i$  и  $\mathbf{q}_i$ ,  $i=1, \dots, N$ ,  $N = N_x N_y$ , для изображений  $D_1^k(x_i, y_j)$  и  $D_2^k(x_i, y_j)$  соответственно. Пусть  $T_1 = \{\mathbf{p}^n\}$ ,  $n=1, \dots, N_1$ , – множество отметок на изображении  $D_1^k(x, y)$ , а  $T_2 = \{\mathbf{q}^m\}$ ,  $m=1, \dots, N_2$ , – множество отметок на изображении  $D_2^k(x, y)$ ;  $N_1, N_2$  – количество отметок на изображениях  $D_1^k(x, y)$  и  $D_2^k(x, y)$  соответственно. Точки  $\mathbf{p} \in D_1^k(x, y)$  и  $\mathbf{q} \in D_2^k(x, y)$  называются сопряженными, если они являются проекциями одной и той же физической точки трехмерной сцены. Известно [3], что для любой точки  $\mathbf{p}$  сопряженная ей точка  $\mathbf{q}$  лежит в изображении  $D_2$  на эпиллярной прямой, определяемой геометрией съемки. Обозначим  $l_1(\mathbf{p})$  эпиллярную прямую в изображении  $D_2^k(x, y)$ , соответствующую точке  $\mathbf{p}$ , и аналогично  $l_2(\mathbf{q})$  – эпиллярную прямую в изображении  $D_1^k(x, y)$ , соответствующую точке  $\mathbf{q}$ . Введем также обозначение  $ld_1(\mathbf{p}) = \{\mathbf{q}_i \mid r(\mathbf{q}_i, l_1(\mathbf{p})) < 0,7\}$ , где  $r(\mathbf{q}_i, l_1(\mathbf{p}))$  – расстояние от точки до прямой. Отметки, соответствующие реальному положению объекта на изображениях  $D_1^k(x, y)$  и  $D_2^k(x, y)$ , очевидно, являются сопряженными точками и, следовательно, должны удовлетворять эпиллярному соотношению. Поэтому если для некоторой отметки из  $T_1$  нет ни одной отметки из  $T_2$ , удовлетворяющей эпиллярному соотношению, то эта отметка ложная. Таким образом, процедура эпиллярного тестирования заключается в построении множеств

$$\hat{T}_1 = \{\mathbf{p}^n \in T_1 \mid ld_1(\mathbf{p}^n) \cap T_2 \neq \emptyset\}, \quad (6)$$

$$\hat{T}_2 = \{\mathbf{q}^m \in T_2 \mid ld_2(\mathbf{q}^m) \cap T_1 \neq \emptyset\}. \quad (7)$$

Вычислим теоретическую вероятность обнаружения при таком способе обработки. Пусть при начальной пороговой обработке величина порога равна  $C_T$ . Тогда вероятность появления отметки в произвольной точке каждого изображения составляет

$$\alpha_1 = 1 - \Phi_{(0,1)}(C_T). \quad (8)$$

Пусть  $\zeta$  – случайная величина, равная количеству ложных обнаруженных объектов на стереопаре. Вычислим математическое ожидание  $E\zeta$ . Рассмотрим произвольную точку  $\mathbf{p}_i$  изображения  $D_1^k(x, y)$ . Если в этой точке сделана отметка, то в зависимости от числа  $k$  отметок на эпилляре  $ld_1(\mathbf{p}_i)$  она порождает  $k$  ложных объектов,  $k=1, \dots, Ne$ , где  $Ne$  – количество точек на эпилляре. Предположим, что  $Ne$  равно для всех эпиллярных прямых. Вероятность того, что  $ld_1(\mathbf{p}_i)$  содержит  $k$  отметок, составляет

$$P(k \text{ отметок на эпилляре}) = C_{Ne}^k \alpha_1^k (1 - \alpha_1)^{Ne - k}. \quad (9)$$

Таким образом, для произвольной точки  $\mathbf{p}_i$  вероятность породить  $k$  ложных объектов составляет

$$P(k \text{ ложных объектов в точке}) = \alpha_1 C_{Ne}^k \alpha_1^k (1 - \alpha_1)^{Ne - k}. \quad (10)$$

Математическое ожидание числа ложных объектов, порождаемых произвольной точкой, равно  $\alpha_1^2 Ne$ . Суммируя по всем точкам изображения, окончательно получим

$$E\zeta = \alpha_1^2 Ne N_x N_y. \quad (11)$$

Поэтому  $\alpha_T = \alpha_1^2 Ne$  является величиной, характеризующей вероятность ложной тревоги при данном методе обработки. Подсчитаем вероятность пропуска объекта. Вероятность того, что на каждом из изображений стереопары в точке, соответствующей реальному положению объекта, не будет сделана отметка, составляет

$$\beta_1 = \Phi_{(0,1)}(C_T - A). \quad (12)$$

Для успешного прохождения эпиполярного теста необходимо, чтобы отметки, соответствующие истинному объекту, были сделаны в обоих кадрах стереопары. Таким образом, вероятность обнаружения равна вероятности совместного обнаружения объекта в каждой монопоследовательности. И окончательно вероятность пропуска объекта для этого алгоритма

$$\beta_T = 1 - (1 - \beta_1)^2. \quad (13)$$

**Алгоритм стереообработки с формированием отметок совместно по двум кадрам.** Вышеизложенная процедура обнаружения объекта, основанная на независимом пороговом обнаружении в каждой последовательности, является неоптимальной. Для повышения вероятности обнаружения необходимо уже на этапе формирования отметок обрабатывать изображения с обеих камер совместно. Заметим, что яркость в точках  $\mathbf{p}_{\text{tar}} = (X_1^k, Y_1^k)$  и  $\mathbf{q}_{\text{tar}} = (X_2^k, Y_2^k)$ , соответствующих реальному положению объекта, имеет распределение  $N(A, 1)$  и соответственно сумма  $D_1^k(\mathbf{p}_{\text{tar}}) + D_2^k(\mathbf{q}_{\text{tar}})$  имеет распределение  $N(2A, 2)$ . Как и раньше, оптимальным будет критерий отношения правдоподобия

$$D_1^k(\mathbf{p}_i) + D_2^k(\mathbf{q}_j) > C_S, \quad (14)$$

где значение порога  $C_S$  определяется вероятностью ложной тревоги, причем рассматриваются только  $\mathbf{p}_i$  и  $\mathbf{q}_j$ , удовлетворяющие эпиполярному соотношению.

Будем называть стереоотметкой тройку чисел, состоящую из пары точек двух изображений, удовлетворяющих эпиполярному соотношению, и суммы яркостей в этих точках:  $s = \langle \mathbf{p}_i, \mathbf{q}_j, D_1(\mathbf{p}_i) + D_2(\mathbf{q}_j) \rangle$ . Построим множество стереоотметок  $S = \{s_i\}$ :

$$S = \bigcup_{i=1}^N \bigcup_{j=1}^N \{ \langle \mathbf{p}_i, \mathbf{q}_j, D_1^k(\mathbf{p}_i) + D_2^k(\mathbf{q}_j) \rangle \mid \mathbf{q}_j \in ld_1(\mathbf{p}_i) \}, \quad (15)$$

а затем проведем его пороговую обработку.

Вычислим  $E\zeta$  для этого алгоритма. Рассмотрим произвольную точку  $\mathbf{p}_i$  изображения  $D_1^k(x, y)$  и обозначим в ней яркость  $\xi$ . Яркость в  $j$ -й точке эпиполяры —  $ld_1(\mathbf{p}_i) - \eta_j$ . Случайные величины  $\xi, \eta_j$  попарно независимы и рас-

пределены по нормальному закону  $N(0,1)$ . Тогда вероятность того, что суммарная яркость точки  $\mathbf{p}_i$  и  $j$ -й точки эпполяры  $ld_1(\mathbf{p}_i)$  больше порога  $C_S$ , равна

$$P(\xi + \eta_j > C_S) = P(\eta_j > C_S - \xi) = 1 - \Phi_{(0,1)}(C_S - \xi) = p(\xi). \quad (16)$$

Следовательно, вероятность того, что суммарная яркость будет больше порогового значения в  $k$  точках, составляет

$$P(k \text{ отметок}) = C_{Ne}^k p^k(\xi)(1 - p(\xi))^{Ne - k}. \quad (17)$$

Отсюда при фиксированном  $\xi$  математическое ожидание ложных объектов, порождаемых точкой первого изображения, равно  $Ne p(\xi)$ . Здесь, как и раньше, будем полагать, что длина эпполяры  $Ne$  равна для всех эпполярных прямых. Используя формулу свертки и суммируя по всем точкам изображения, получаем математическое ожидание числа ложных объектов, обнаруживаемых по стереопаре:

$$E\xi = N_x N_y \int_{-\infty}^{\infty} Ne(1 - \Phi_{(0,1)}(C_S - x))\Phi_{(0,1)}(x) dx. \quad (18)$$

Поэтому

$$\alpha_S = \int_{-\infty}^{\infty} Ne(1 - \Phi_{(0,1)}(C_S - x))\Phi_{(0,1)}(x) dx \quad (19)$$

является величиной, характеризующей вероятность ложной тревоги при данном методе обработки. Поскольку сумма яркостей в стереоотметке, соответствующей реальному объекту, имеет распределение  $N(2A, 2)$ , то вероятность пропуска объекта

$$\beta_S = \Phi_{(2A, 2)}(C_S) = \Phi_{(0,1)}(C_S/\sqrt{2} - A\sqrt{2}). \quad (20)$$

Оба алгоритма стереобнаружения позволяют в отличие от моноалгоритма определить трехмерные координаты объекта.

На рис. 1,  $a, b$  изображены графики зависимости вероятности обнаружения от уровня ложной тревоги для рассматриваемых алгоритмов при  $A = 4$  ( $a$ ) и  $A = 5$  ( $b$ ) соответственно ( $\blacktriangle$  – монообработка,  $\bullet$  – триангуляционный алгоритм,  $\blacksquare$  – стереообработка совместно по двум кадрам). Значения  $N_x, N_y$  равны 256, в качестве оценки длины эпполяры  $Ne$  также использовалось значение 256.

**Экспериментальные результаты.** Для исследования эффективности предложенных алгоритмов был смоделирован процесс наблюдения за движущимся объектом со спутников, летящих по круговой орбите. При генерации изображений использовался программный продукт Pov-Ray версии 3.1. В течение всего времени наблюдения каждая камера отслеживала фиксированную точку на подстилающей поверхности и наблюдаемый объект находился в поле зрения обеих камер. Детальное исследование различных процедур межкадровой обработки для рассматриваемого случая проведено в работе [4]. Исходный шум добавлялся таким образом, чтобы после МКО соотно-

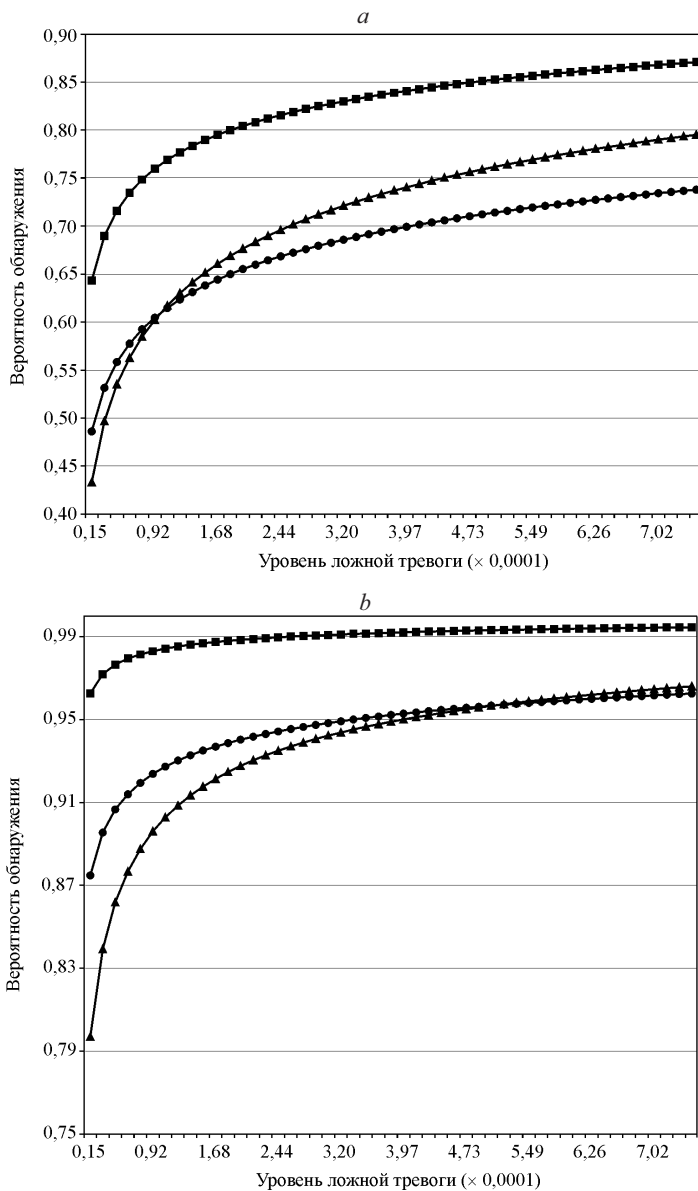


Рис. 1

шение исходной яркости объекта и СКО шума было равно 4 и 5 соответственно. Среднее значение длины эпиполяры для использованной геометрии съемки было порядка 250 пикселей. На рис. 2, *a*, *b* показаны экспериментально оцененные зависимости вероятности обнаружения от уровня ложных тревог. (Обозначения на рисунке такие же, как на рис. 1.) Из графиков видно, что алгоритмы, использующие стереоданные в процессе формирования отметок, позволяют с большей вероятностью обнаруживать объекты. Алгоритм триангуляции имеет лучшие по сравнению с монообработкой характеристики при обнаружении объектов с яркостью 5 и более СКО шума и позволяет определить трехмерные координаты объекта.

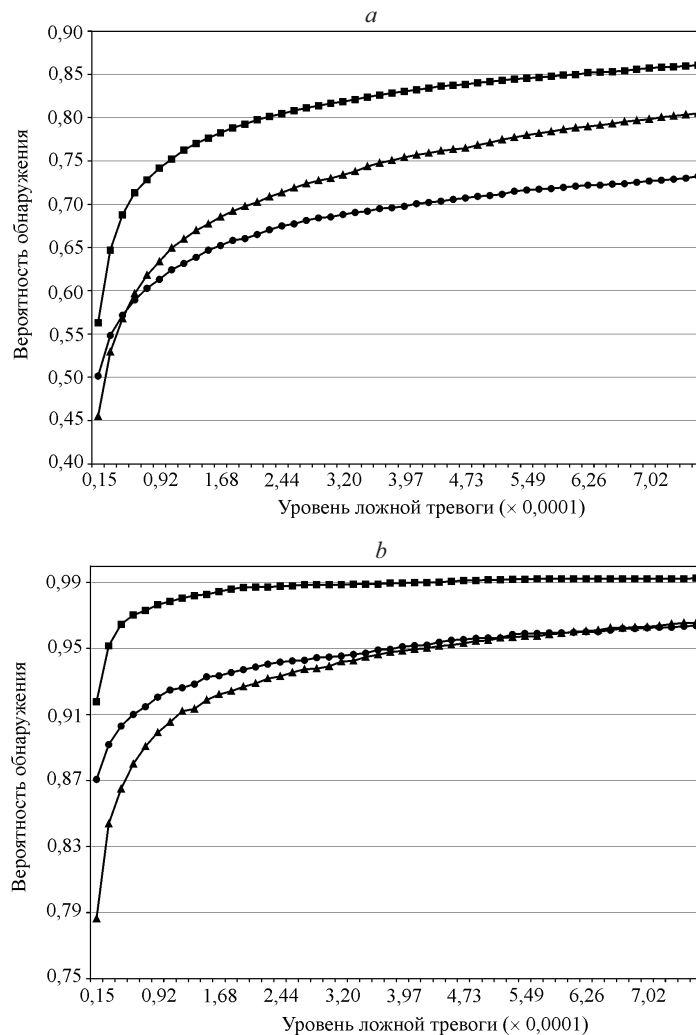


Рис. 2

На практике построение более эффективных алгоритмов обработки продиктовано стремлением обнаруживать менее яркие объекты при фиксированных значениях вероятности обнаружения и ложной тревоги. Из приведенных выше формул легко получить зависимость амплитуды от уровня ложной тревоги при фиксированной вероятности обнаружения. На рис. 3 представлен теоретически полученный график зависимости амплитуды объекта от уровня ложной тревоги при вероятности обнаружения, равной 0,9. (Обозначения кривых на рисунке такие же, как на рис. 1.)

Важной особенностью стереонаблюдения является возможность использования априорной информации о допустимой высоте положения объекта. При стереонаблюдении для полученных отметок можно восстановить трехмерные координаты и исключить из рассмотрения точки, лежащие вне заданного диапазона высот. Для исследования эффективности информации о высоте объекта нами был проведен следующий эксперимент. Диапазон высот, содержащий объект, выбирался так, что полоса неопределенности вдоль

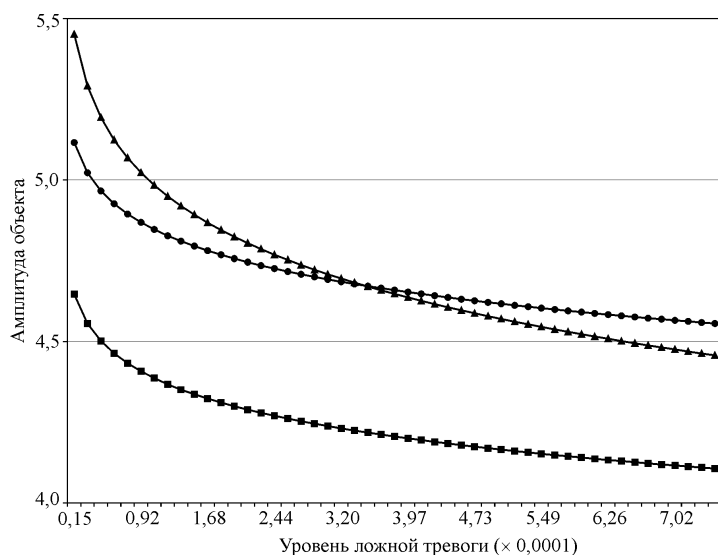


Рис. 3

$N_e$	Вероятность ложной тревоги		
	Монообработка	Триангуляционный алгоритм	Стереобработка совместно по двум кадрам
16	$2,63 \cdot 10^{-3}$	$0,99 \cdot 10^{-3}$	$0,45 \cdot 10^{-4}$
32	$2,63 \cdot 10^{-3}$	$2,02 \cdot 10^{-3}$	$0,91 \cdot 10^{-4}$
64	$2,63 \cdot 10^{-3}$	$4,07 \cdot 10^{-3}$	$1,67 \cdot 10^{-4}$
128	$2,63 \cdot 10^{-3}$	$8,10 \cdot 10^{-3}$	$3,05 \cdot 10^{-4}$

эпиполярной прямой составляла заданное число пикселей  $N_e$ . Затем, считая высоты априорно известной информацией, проводилось обнаружение объекта по уже предложенной схеме. Амплитуда объекта в данном эксперименте составляла 4 СКО шума, вероятность обнаружения 0,9. Полученные вероятности ложной тревоги в зависимости от количества пикселей на эпиполяре для рассматриваемых алгоритмов приведены в таблице.

**Заключение.** Теоретические и экспериментальные исследования показали целесообразность применения стереовидения в задаче обнаружения малоразмерных динамических объектов не только с точки зрения получения информации о трехмерном положении, но и с точки зрения повышения вероятности их обнаружения. Установлено, что формирование множества отметок отдельно по каждой монопоследовательности является неоптимальным, однако позволяет получить существенный выигрыш по сравнению с монообнаружением при наблюдении относительно ярких целей. В то же время, используя всю совокупность стереоданных в процессе формирования отметок, можно понизить уровень ложной тревоги приблизительно в 10 раз.



Экспериментально подтверждена возможность эффективного учета информации о высотном ходе объекта при любом из рассмотренных способов стереообнаружения.

#### СПИСОК ЛИТЕРАТУРЫ

1. **Киричук В. С., Иванов В. А., Ангеров В. Ю., Синельщиков В. В.** Методика выделения подвижных точечных объектов при регистрации сцены с двух точек наблюдения // Автометрия. 2000. № 6. С. 3.
2. **Боровков А. А.** Математическая статистика. Новосибирск: Наука, 1997. С. 342.
3. **Faugeras О.** Three-Dimensional Computer Vision. London: MIT Press, 1999. P. 169.
4. **Киричук В. С., Шакенов А. К.** Обнаружение точечных динамических объектов, регистрируемых движущейся камерой // Автометрия. 2004. **40**, № 1. С. 3.

*Институт автоматизации и электрометрии СО РАН,  
E-mail: adil.shakenov@ngs.ru*

*Поступила в редакцию  
4 июня 2004 г.*