

УДК 004.392

ИЗМЕРЕНИЕ КООРДИНАТ ОБЪЕКТА В ПОСЛЕДОВАТЕЛЬНОСТИ МУЛЬТИСПЕКТРАЛЬНЫХ ИЗОБРАЖЕНИЙ

Алпатов Б.А., д.т.н., проф., зав. каф. Автоматики и информационных технологий в управлении Рязанского государственного радиотехнического университета (РГРТУ), e-mail: aitu@rsreu.ru;

Бабаян П.В., к.т.н., доцент, начальник лаборатории РГРТУ, e-mail: aitu@rsreu.ru;

Смирнов С.А., аспирант РГРТУ, e-mail: aitu@rsreu.ru;

Фельдман А.Б., аспирант РГРТУ, e-mail: aitu@rsreu.ru.

Ключевые слова последовательность изображений, мультиспектральное наблюдение, измерение координат, оптико-электронные системы.

Введение

Одним из важных практических приложений цифровой обработки изображений является разработка бортовых оптико-электронных систем. Основным классом задач, решаемых в таких системах, является обнаружения и измерение координат объектов интереса. Высокая степень априорной неопределенности относительно условий наблюдения и необходимость работы в реальном масштабе времени накладывают жесткие ограничения на используемые алгоритмы анализа изображений.

Бортовые оптико-электронные системы, как правило, должны функционировать в сложных погодных условиях, таких как дождь, туман или облачность, и в любое время суток, в том числе ночью. Одним из способов повышения эффективности применения бортовых систем является использование мультиспектральных датчиков изображений. Конструктивно такой датчик обычно представляет собой пару оптически съюстированных датчиков телевизионного (ТВ) и инфракрасного (ИК) диапазонов. При этом возникает проблема совместного использования информации, получаемой от датчиков, что предполагает создание алгоритмов анализа мультиспектральных изображений. Данная проблема тесно связана с задачами комплексирования изображений.

На практике принято выделять три уровня комплексирования [1]. Первый уровень предполагает, что изображения двух каналов объединяются с целью получения нового улучшенного изображения [2]. Результирующее изображение в дальнейшем используется для решения конкретных прикладных задач. Второй уровень связан с объединением признаков. При этом признаки выделяются отдельно на каждом изображении и затем объединяются по заданному правилу, после чего решаются задачи обнаружения и измерения координат объекта интереса [3, 4]. Наконец, третий уровень предполагает объединение конечных результатов независимой обработки изображений каждого спектрального канала. Под этим может подразумеваться, например, принятие решения об обнаружении объекта на основе решений, принятых по каждому изображению в отдель-

Рассматривается ряд разработанных авторами алгоритмов измерения координат объекта в последовательности мультиспектральных изображений. В ходе экспериментов, выполненных с использованием натуральных видеосюжетов, показано, что совместное использование информации ТВ и ИК диапазонов позволяет повысить эффективность слежения за объектом. Показано, что дополнительного улучшения можно добиться, если на этапе измерения координат учитывать неоднородность фонового изображения и наличие посторонних мешающих объектов.

ности, или уточнение координат объекта на основе совместной обработки независимых измерений в каждом канале [5-7].

Надо заметить, что алгоритмы первого уровня комплексирования изображений применяются, как правило, для улучшения восприятия человеком фоноцелевой обстановки [2] и, лишь в значительно меньшей степени, в задачах автоматического анализа изображений. Пример использования второго уровня комплексирования приведен, например, в статье [3]. В каждом канале независимо выполняется выделение движущегося объекта, после чего полученные бинарные изображения объединяются, образуя результирующее множество точек объекта. Использование третьего уровня комплексирования возможно, если задана или может быть оценена степень доверия к результатам, полученным в каждом отдельном канале [7].

В данной работе рассматривается три алгоритма измерения координат. Первый согласно приведенной выше классификации относится к третьему уровню комплексирования. Два других алгоритма можно отнести ко второму уровню комплексирования. Все три алгоритма основываются на принципе сопоставления текущего изображения с эталонным изображением объекта, получившим на практике самое широкое распространение благодаря своей надежности и помехоустойчивости.

Сходство фрагмента наблюдаемого изображения с эталоном часто определяют при помощи критерия [8]

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

где (α, β) – координаты анализируемого фрагмента текущего изображения $l(i, j, n)$, n – номер кадра, $h(i, j, n)$ – эталонное изображение, заданное на множестве $H(n)$. Для определения координат объекта критерий (1) необходимо минимизировать:

$$(\alpha^*(n), \beta^*(n)) = \arg \min_{(\alpha, \beta)} F(\alpha, \beta, n). \quad (2)$$

С целью учета эффектов, связанных с изменчивостью объекта, вводится дополнительный шаг, который заключается в рекурсивной фильтрации эталона [9]:

$$h(i, j, n) = \gamma h(i, j, n-1) + (1-\gamma)l(i + \alpha^*(n), j + \beta^*(n), n), \quad (3)$$

где $\gamma \in [0..1]$ – коэффициент, определяющий эффективную память фильтра. Назовем (1)-(3) базовым алгоритмом.

Рассмотренные в статье алгоритмы можно кратко охарактеризовать следующим образом. В алгоритме №1 операции (1) и (2) выполняются для каждого спектрального канала, после чего осуществляется выбор наиболее подходящего из них в соответствии с подходом, описанным в работе [7]. Алгоритм №2 представляет собой модификацию базового алгоритма, позволяющую перейти к обработке мультиспектральных изображений. Алгоритм №3 позволяет дополнительно учесть присутствие в кадре похожих на объект интереса фоновых образований и наличие посторонних мешающих объектов. Приведем краткое описание перечисленных алгоритмов.

Алгоритм №1. Данный алгоритм основывается на том факте, что в отдельные моменты времени один из спектральных каналов по отношению к другому может иметь лучшие характеристики наблюдения объекта интереса. Критерий, по которому следует выбирать наиболее подходящий канал, очевидно, должен быть связан с используемым алгоритмом измерения координат, причем оценка численного значения критерия должна выполняться апостериори, т.е. после того, как сформированы предварительные измерения в каждом из каналов.

В настоящей работе для измерения координат объекта в каждом канале будем использовать алгоритм, описываемый выражениями (1)-(2). После формирования предварительных измерений в каждом канале $(\hat{\alpha}_k(n), \hat{\beta}_k(n))$, где $k = 1, 2$ – номер спектрального канала, для каждого из них следует вычислить показатель работоспособности [7]:

$$D_k(n) = \frac{\min_{(\alpha, \beta) \in \Gamma} F(\alpha, \beta, n-1)}{\min_{(\alpha, \beta)} F(\alpha, \beta, n)}, \quad k = 1, 2, \quad (4)$$

где $\min_{(\alpha, \beta)} F(\alpha, \beta, n)$ следует интерпретировать, как оценку степени временной изменчивости изображения объекта; $\min_{(\alpha, \beta) \in \Gamma} F(\alpha, \beta, n-1)$, в свою очередь, дает оценку отличия изображения объекта от окружающего фона; Γ – множество граничных точек зоны поиска.

Показатель (4) позволяет выбрать наиболее подходящий канал и полученные в нем измерения координат объекта:

$$(\alpha^*(n), \beta^*(n)) = \begin{cases} (\hat{\alpha}_1(n), \hat{\beta}_1(n)), & D_1(n) \geq D_2(n); \\ (\hat{\alpha}_2(n), \hat{\beta}_2(n)), & \text{иначе.} \end{cases} \quad (5)$$

Выборное таким образом измерение используется

для коррекции текущих параметров слежения в канале, имеющем меньший показатель работоспособности. При этом обновление изображений объекта в каждом канале выполняется согласно выражению (3) с учетом (5).

Алгоритм №2. Запишем модель формирования мультиспектрального изображения при отсутствии фона

$$\mathbf{l}(i, j, n) = r(i, j, n)\mathbf{h}(i, j, n) + \xi(i, j, n),$$

$$\mathbf{l}(i, j, n) = [l_k(i, j, n)]^T, \quad \mathbf{h}(i, j, n) = [h_k(i, j, n)]^T,$$

$$\xi(i, j, n) = [\xi_k(i, j, n)]^T, \quad k = 1, 2. \quad (6)$$

Функции $l_k(i, j, n)$, $h_k(i, j, n)$, $\xi_k(i, j, n)$ задают соответственно наблюдаемое изображение, эталонное изображение объекта и шумовой процесс в k -м спектральном канале; $r(i, j, n)$ – индикаторная функция, значения которой равны 1 в точках, принадлежащих объекту, и 0 в противном случае. Значения $\xi_k(i, j, n)$ некоррелированы по i, j, n и k , имеют нормальное распределение с нулевым средним и дисперсией $\sigma_k^2(n)$.

Для модели (5) оптимальный по критерию максимального правдоподобия алгоритм измерения координат объекта связан минимизацией функции

$$F_1(\alpha, \beta, n) = \sum_{k=1}^2 \frac{1}{\sigma_k^2(n)} \sum_{(i, j) \in H(n)} (l_k(i + \alpha, j + \beta, n) - h_k(i, j, n))^2. \quad (7)$$

Эксперименты показывают, что корреляция яркостей точек объекта между каналами оказывается незначительной, поэтому оценки $\hat{h}_k(i, j, n)$ изменяющихся изображений объекта будем формировать независимо в каждом канале в соответствии с выражением (3).

На практике интенсивность шума в датчиках изображений достаточно мала, однако слагаемое $\xi(i, j, n)$ в модели (6) позволяет учесть некоторые эффекты, которые сами по себе довольно трудно описать аналитической моделью. Например, для оценки яркостей точек объекта требуется некоторое время. Между тем объект постоянно меняется вследствие изменения ракурса и условий наблюдения. Таким образом, шумовые процессы $\xi_k(i, j, n)$, $k = 1, 2$ в модели (6) описывают фактически имеющее место отличие наблюдаемого изображения объекта от его эталона. С этой точки зрения предпочтительным является тот канал, где это отличие меньше. Поэтому, исходя из предположения, что к предшествующему кадру получена достаточно хорошая оценка эталонного изображения, координаты объекта $\alpha^*(n-1)$, $\beta^*(n-1)$ были измерены верно и изменчивость изображений за период формирования очередного кадра незначительна, оценку дисперсии шума будем вычислять согласно выражению

$$\hat{\sigma}_k^2(n+1) = \frac{1}{K_H(n)} \sum_{(i, j) \in H(n)} (l_k(i + \alpha^*(n), j + \beta^*(n), n) - h_k(i, j, n))^2, \quad (8)$$

$$k = 1, 2$$

где $K_H(n)$ – мощность множества $H(n)$.

Согласно (7) критерий сходства изображений вычисляется независимо в каждом канале, после чего, после чего результаты суммируются с весами, зависящими от дисперсий шума. Таким образом, алгоритм №2 может быть отнесен ко второму уровню комплексирования.

Алгоритм №3. Модель (6) предполагает отсутствие фона и посторонних объектов. Между тем именно наличие похожих на объект интереса фоновых образований и мешающих объектов является одной из основных причин возникновения аномальных ошибок измерения координат. Борьба с этим можно путем уменьшения размера зоны поиска, что позволяет, кроме того, сократить вычислительные затраты. Однако сильное уменьшение зоны поиска препятствует эффективному отслеживанию маневров объекта, повышает вероятность срыва слежения из-за неконтролируемого движения носителя и снижает вероятность повторного захвата при возникновении аномальных ошибок измерения координат.

Чтобы избежать появления аномальных ошибок, не прибегая к существенному уменьшению зоны поиска, предлагается учитывать неоднородность фона и наличие посторонних движущихся объектов непосредственно на этапе измерения координат объекта интереса [10]. Рассмотрим модель наблюдения

$$\mathbf{l}(i, j, n) = r(i, j, n)\mathbf{h}(i, j, n) + (1 - r(i, j, n))\mathbf{q}(i, j, n) + \xi(i, j, n), \quad (9)$$

где $\mathbf{q}(i, j, n) = [q_k(i, j, n)]^T$, $k = 1, 2$ – изображение, представляющее собой изображение фона $\mathbf{g}(i, j, n) = [g_k(i, j, n)]^T$ с наложенными изображениями посторонних объектов; считаем, как и ранее, что изображения в разных каналах не коррелированы между собой. Для простоты дальнейшего изложения будем полагать, что фон неподвижен, т.е. между следующими подряд кадрами отсутствуют геометрические преобразования. В противном случае геометрические преобразования могут быть скомпенсированы на основании известных подходов [11].

Если шумовые процессы $\xi_k(i, j, n)$, $k = 1, 2$ являются гауссовыми с нулевым средним и дисперсией $\sigma_k^2(i, j, n)$ в каждой точке, то в соответствии с критерием максимального правдоподобия необходимо минимизировать критерий

$$F_2(\alpha, \beta, n) = \sum_{k=1}^2 \sum_{(i,j) \in H(n)} \frac{(l_k(i+\alpha, j+\beta, n) - h_k(i, j, n))^2 - (l_k(i+\alpha, j+\beta, n) - q_k(i, j, n))^2}{\sigma_k^2(i, j, n)} \quad (10)$$

При наличии похожих на объект интереса посторонних объектов или участков фона второе слагаемое в числителе позволит избежать возникновения дополнительных локальных минимумов.

В выражении (10) неизвестными являются изображения $q_k(i, j, n)$ и $h_k(i, j, n)$, а также дисперсии $\sigma_k^2(i, j, n)$. Изображения $q_k(i, j, n)$, $k = 1, 2$ при этом представляют собой композицию изображений фона

$g_k(i, j, n)$ и неизвестных априори изображений посторонних движущихся объектов.

Оценки изображений объекта интереса и фона найдем посредством рекурсивной фильтрации:

$$\begin{aligned} \hat{h}_k(i, j, n) &= \hat{h}_k(i, j, n-1)\gamma + \\ &+ l_k \left[(i + \alpha^*(n), j + \beta^*(n), n) \right] (1 - \gamma), (i, j) \in H(n) \\ \hat{g}_k(u, v, n) &= \begin{cases} \hat{g}_k(u, v, n-1)\gamma + l_k(u, v, n)(1 - \gamma), (u, v) \in G(n); \\ \hat{g}_k(u, v, n-1), \text{ иначе.} \end{cases} \quad (11) \end{aligned}$$

Здесь $G(n)$ – множество точек фона, не закрытых в объектами в n -м кадре.

Оценивать дисперсию шума в данном случае целесообразно совместно с оцениванием фона. На начальном этапе, пока не получено достоверных оценок фона, при помощи фильтра с растущей памятью вычисляется второй начальный момент

$$\begin{aligned} \hat{\mu}_k^2(u, v, n) &= \frac{n-1}{n} \hat{\mu}_k^2(u, v, n-1) + \\ &+ \frac{1}{n} l_k^2(u, v, n), (u, v) \in G(n), n < N_0, k = 1, 2. \end{aligned} \quad (12)$$

При достижении момента времени $n = N_0$, можно перейти к оценке дисперсии

$$\hat{\sigma}_k^2(u, v, N_0) = \hat{\mu}_k^2(u, v, N_0) - g_k^2(u, v, N_0), k = 1, 2. \quad (13)$$

Дальнейшая оценка выполняется в соответствии с выражением

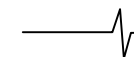
$$\begin{aligned} \hat{\sigma}_k^2(u, v, n) &= \gamma \hat{\sigma}_k^2(u, v, n-1) + \\ &+ (1 - \gamma) (l_k(u, v, n) - g_k(u, v, n))^2, \\ (u, v) \in G(n), n > N_0, k = 1, 2. \end{aligned} \quad (14)$$

Для формирования оценки $\hat{q}_k(i, j, n)$ будем вместо неизвестных изображений посторонних объектов брать соответствующие им фрагменты наблюдаемого изображения, задаваемые индикаторной функцией $b(i, j, n)$. Функция $b(i, j, n)$ формируется по следующему правилу:

$$b(i, j, n) = \begin{cases} 1, & |l_k(i, j, n) - \hat{g}_k(i, j, n)| > t \hat{\sigma}_k^2(i, j, n), k = 1, 2; \\ 0, & \text{иначе.} \end{cases} \quad (15)$$

Операция (15) позволяет выделить изменяющиеся фрагменты наблюдаемого изображения, которые, обычно, соответствуют движущимся объектам, при этом коэффициент t определяет допустимый уровень ложных тревог.

Если объект интереса движется, то необходима дополнительная обработка индикаторной функции $b(i, j, n)$. Для этого вначале выполняются морфологические операции [12], после чего полученные связные области анализируются на предмет их близости к положению объекта интереса на предыдущем кадре. Предпола-



гается, что ближайший сегмент принадлежит объекту, и соответствующие ему точки $b(i, j, n)$ устанавливаются в нулевое значение. Результатом этих операций будет модифицированная индикаторная функция, которую обозначим $\tilde{b}(i, j, n)$. При этом окончательное выражение для формирования оценки $\hat{q}_k(i, j, n)$ принимает вид:

$$\hat{q}_k(i, j, n) = \tilde{b}(i, j, n)l_k(i, j, n) + (1 - \tilde{b}(i, j, n))\hat{g}_k(i, j, n), \quad k = 1, 2. \quad (16)$$

Формула (10) показывает, что вычисления можно организовать параллельно в каждом спектральном канале, после чего следует объединить полученные результаты, которые в дальнейшем используются для оценки положения объекта в кадре. Таким образом, организация алгоритма №3 соответствует второму уровню комплексирования.

Экспериментальные исследования

Исследования проводились на базе из 29 натуральных мультиспектральных видеосюжетов. Длительность сюжетов варьировалась от 100 до 400 кадров, средняя длительность – 229 кадров. Объекты интереса – движущиеся по трассе автомобили, наблюдаемые на удалении от 1000 до 2000 м. Размер изображений объектов интереса варьировался от 14x10 до 43x22 пикселей, средний размер 23x14 пикселей. На видеосюжетах присутствовали похожие на объект интереса движущиеся объекты (другие автомобили), имело место частичное заслонение объектов интереса элементами фона (кустарник, строения). Человеком-экспертом указывалось эталонное (истинное) положение объекта интереса в кадре. При возникновении аномальных ошибок по истинным координатам формировалось новое целеуказание, на основе которого выполнялось обновление эталонных изображений. Примеры кадров используемых видеосюжетов представлены на рис. 1.

Для описания результатов экспериментов введем ряд определений.

1) *Аномальная ошибка* – непопадание измеренных координат, в прямоугольник, ограничивающий изображение объекта в кадре.

2) *Слежение* – измерение координат объекта без аномальных ошибок на протяжении не менее 5 следующих подряд кадров.

3) *Срыв* – прерывание слежения.

Алгоритмы измерения координат, с точки зрения достигаемой при их использовании эффективности слежения за объектом характеризуют следующие показатели.

1) *Вероятность правильного сопоставления* (P_D) – отношение числа кадров без аномальных ошибок к общей длительности видеосюжета.

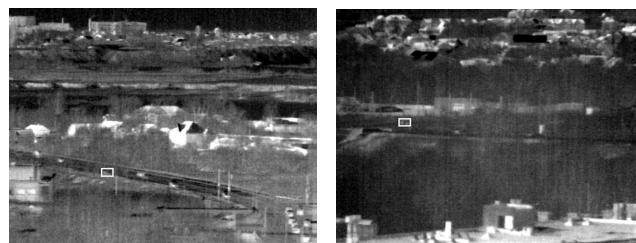
2) *Средняя продолжительность слежения* (N_{cp}).

3) *Отношение длительности периодов слежения* N_{cp} к длине видеосюжета (f_{cp})

4) *Вероятность срыва* (P_A).



а)



б)

Рис. 1 Примеры кадров из тестовых видеосюжетов: а) – ТВ канал; б) – ИК канал (объект показан в рамке)

В ходе исследований изображения дополнительно искажались аддитивным гауссовским шумом со среднеквадратическим отклонением $\sigma = 4$. Размер зоны поиска – 51x51 пикселей. Показатели эффективности вычислялись на основе обработки каждого видеосюжета, затем усреднялись на всей выборке видеосюжетов с учетом их длительности.

Сравнительные данные по эффективности исследуемых алгоритмов приведены в табл 1.

Таблица 1. Сравнение эффективности исследуемых алгоритмов (аддитивный шум $\sigma = 4$, зона поиска – 51x51 пикселей)

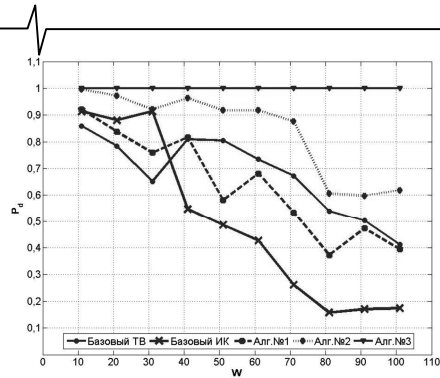
	Базовый алгоритм (ТВ)	Базовый алгоритм (ИК)	Алг. №1	Алг. №2	Алг. №3
P_D	0,849	0,750	0,895	0,909	0,978
N_{cp}	111,20	82,22	131,97	157,95	189,90
f_{cp}	0,487	0,399	0,592	0,666	0,779
P_A	0,008	0,010	0,007	0,005	0,005

На рис. 2 изображены графики зависимостей вероятности правильного сопоставления от размера зоны поиска для видеосюжетов, показанных на рис. 1. Графики показывают, что только для алгоритма №3 увеличение зоны поиска объекта интереса не приводит к росту числа аномальных ошибок.

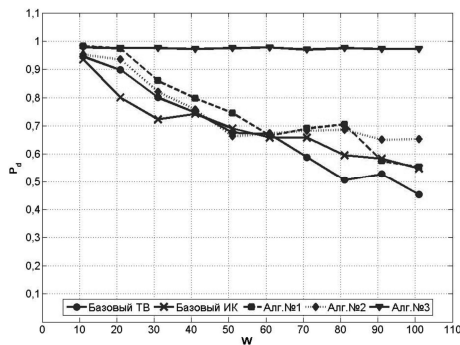
Заключение

Совместное использование датчиков изображений разного спектрального диапазона позволяет уменьшить число срывов и увеличить длительность слежения за объектом. В статье описаны возможные подходы к анализу мультиспектральных видеопоследовательностей.

Эффективность этих подходов в сравнении с независимой обработкой изображений в отдельных спектральных каналах продемонстрирована в ходе экспериментальных исследований, выполненных на большой базе натуральных видеосюжетов.



а)



б)

Рис. 2 Графики зависимостей вероятности правильного сопоставления от размера зоны поиска W для сюжетов, показанных на рис. 1 (СКО аддитивного шума $\sigma = 6$)

Результаты исследований показывают, что большего повышения эффективности можно добиться, если дополнительно парировать негативные эффекты, связанные с присутствием похожих на объект интереса фоновых образований и посторонних движущихся объектов. Фон и мешающие объекты непосредственно учитываются на этапе измерения координат, что позволяет снизить вероятность появления аномальных ошибок без существенного уменьшения зоны поиска объекта.

Исследования выполнены при поддержке Министерства образования и науки РФ (государственный контракт № 14.740.11.0284)

Литература

1. J. Dong, D. Zhuang, Y. Huang and J. Fu. Advances in Multi-Sensor Data Fusion: Algorithms and Applications // Sensors. – 2009. – №9(10). p. 7771-7784.
2. Lanir J. Maltz M., Rotman S.R. Comparing multispectral image fusion methods for a target detection task // Optical Engineering. – 2007. – Vol. 46(6). – P. 066402-1–066402-8
3. 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.

4. Patrick J. Bonnin Parallel cooperative segmentation method for multispectral images // Visual Information Processing IV Proc. SPIE, Vol. 2488, 207 (1995)

5. Vidya Manian, Luis O. Jimenez, Land cover and ben-thic habitat classification using texture features from hyper-spectral and multispectral images // Journal of Electronic Imaging 16(2), 023011 (Apr–Jun 2007), pp. 1-12.

6. Laurent Nolibe, Julien Borgnino, Marc Ducoulombier, Michel Artaud, Adaptive multispectral detection of small targets using spatial and spectral convergence factor // Proc. of SPIE. Vol. 2759, Signal and Data Processing of Small Targets 1996, pp. 111-120.

7. Бабаян П.В., Смирнов С.А. Слежение за объектом на основе алгоритма сопоставления с эталоном при одновременном наблюдении в видимом и инфракрасном диапазонах // Цифровая обработка сигналов. – №4. – 2010. – С.18-21.

8. 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.

9. Алпатов Б. А., Селяев А. А. Алгоритм оценки местоположения объекта на двумерном изображении. // Изв. вузов. Сер. Приборостроение. – 1988. – №5. – С. 3-5.

10. Бабаян П.В., Фельдман А.Б. Определение местоположения объекта на изображении в системах технического зрения мобильных робототехнических комплексов // Вестник Рязанского государственного радиотехнического университета. – Рязань, 2011. – №4. – С.19-25.

11. Алпатов Б.А., Стротов В.В. Алгоритм оценки параметров геометрических преобразований изображений при мультиспектральном наблюдении // Цифровая обработка сигналов. – 2010. – №4. – С.2-6.

12. Гонсалес Р., Вудс Р. Цифровая обработка изображений. – М.: Техносфера, 2005 г. – 1072 с.

THE OBJECT POSITION MEASUREMENT ON MULTISPECTRAL IMAGES

Alpatov B.A., Babayan P.V., Smirnov S.A., Feldman A.B.

In this paper several algorithms of the object position measurement on multispectral images is considered. The experiments show that combined use of images from TV and IR bands improves the efficiency of object tracking. For preventing problems which may appear because of presence of non-uniform background and other moving objects, it is proposed to take into account these effects directly on the stage of the measurement of the object position.