

УДК 004.932.2

## ОСОБЕННОСТИ РАБОТЫ АЛГОРИТМОВ ОБНАРУЖЕНИЯ УПАВШЕГО ЧЕЛОВЕКА ПО ПОСЛЕДОВАТЕЛЬНОСТИ ИЗОБРАЖЕНИЙ СЦЕНЫ

В. А. Иванов, В. С. Киричук

*Учреждение Российской академии наук  
Институт автоматизи и электрометрии Сибирского отделения РАН,  
630090, г. Новосибирск, просп. Академика Коптюга, 1  
E-mail: vaivanov@iae.nsk.su*

Дан анализ особенностей и предложены подходы к повышению устойчивости работы алгоритмов обнаружения упавшего человека по последовательности изображений, получаемых стационарными камерами видеонаблюдения, при невысокой вычислительной сложности. Рассмотрен вариант получения более достоверного силуэта объекта, проведён анализ и отбор параметров, используемых при распознавании состояния объектов. Представлена методика повышения устойчивости детектора, заключающаяся в выделении интервалов стабильности средних значений параметров, что позволяет повысить достоверность идентификации состояния объектов. Приведены результаты вычислительных экспериментов.

*Ключевые слова:* объект, последовательность, изображение, фон, детектор, распознавание.

**Введение.** Несвоевременное обнаружение упавшего человека, получившего повреждения, приводит к усугублению его состояния. Контроль за пожилыми и больными людьми может осуществляться индивидуальными приборами на основе акселерометров и гироскопов, которые формируют сигнал тревоги при падении человека [1]. Задача обнаружения упавшего человека с помощью систем видеонаблюдения является важной в частном жилье, медицинских учреждениях, местах повышенной опасности.

Имеется обширная литература, посвящённая данной тематике [1–4], однако рассматриваемые в ней подходы и алгоритмы не решают в полной мере эту задачу.

В представленной работе проведён анализ особенностей и предложены подходы к повышению устойчивости работы алгоритмов обнаружения упавшего человека по последовательности изображений, получаемой стационарной камерой наблюдения, при невысокой вычислительной сложности.

**Особенности алгоритмов.** Из анализа публикаций следует, что к настоящему времени сложилась структура алгоритмов обнаружения упавших объектов в видеопотоке (рис. 1).

Выделение движущихся объектов построено по стандартной схеме, использующей оцениваемый фон. В [5–8] широко представлены различные алгоритмы оценивания фона и его обновления, которые в целом удовлетворительно решают задачу разделения сцены на два множества.

Для выделения движущегося объекта в текущем кадре  $t$  вычисляется разность:

$$D_t(x, y) = \left| B_t(x, y) - I_t(x, y) \right|. \quad (1)$$

Здесь  $B_t(x, y)$  — фон;  $I_t(x, y)$  — текущий кадр последовательности;  $(x, y)$  — координаты пикселей фона и текущего изображения ( $x = \overline{1, W}$ ,  $y = \overline{1, H}$ ;  $W$  и  $H$  — размеры кадра

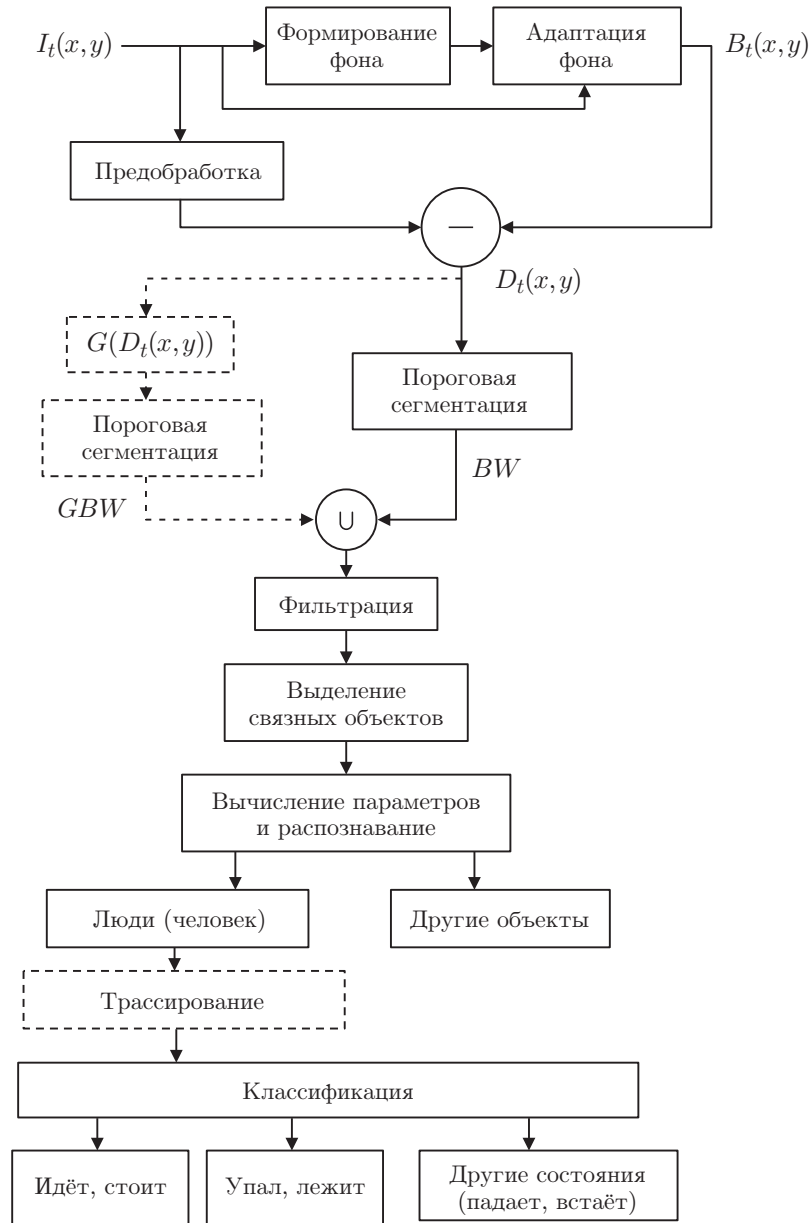


Рис. 1

по  $(x, y)$ ). Для модификации фона обычно используется оценка скользящего среднего на основе соотношения

$$B_t(x, y) = (1 - \alpha)B_{t-1}(x, y) + \alpha I_t(x, y), \quad (2)$$

где  $1/\alpha$  — «время обучения» или время запаздывания оценки фона. Маскирование областей с движущимися объектами в (2) применяется для уменьшения искажения фона при его модификации [6]. Другой подход к оцениванию фона предложен в [8]. Выделение объектов проводится с помощью пороговой сегментации, разделяющей сцену на фон и объекты, подлежащие дальнейшему анализу:

$$BW = \begin{cases} 1, & D(x, y) > Pr, \\ 0, & D(x, y) \leq Pr, \end{cases} \quad (3)$$

где  $P_r$  — порог.

Качество получаемого бинарного изображения  $BW$  является определяющим при дальнейшем его использовании. Существуют три обстоятельства, которые значительно искажают силуэт объектов и влияют на дальнейшую обработку, а именно: движущиеся с объектом тени в текущем кадре; близкие значения яркости объектов на текущем кадре и фона (особенно для серых изображений); положение человека относительно камеры.

Имеющиеся подходы для выделения и компенсации теней в основном работают с цветными изображениями и слабоконтрастными тенями [9–11]. Алгоритмы не работают на тёмных тенях и сценах со снежным покровом. Задача удовлетворительно может быть решена при наличии практически полной информации о рельефе сцены и положении источника света. Для серых изображений задача значительно сложнее и решается в немногих частных случаях.

Совпадение яркости областей объекта и фона также искажает силуэт объекта и тем самым вносит значительные ошибки в оцениваемые по бинарному изображению параметры, что приводит к неверной классификации состояния объекта.

На рис. 2, *a* представлен фон (кадр без объекта), на рис. 2, *b* — кадр с объектом, на рис. 2, *c* приведена разность  $D(x, y)$ , а на рис. 2, *d* — бинарное изображение  $BW$ , полученное из  $D(x, y)$  согласно (3). Из рисунка видно, что яркость решётки фона совпадает с яркостью куртки человека, и это приводит к разделению связного объекта, а далее — к неверным оценкам параметров. Для заполнения утраченных в  $BW$  областей предлагается подвергнуть разность  $D(x, y)$  преобразованию  $G(D(x, y))$ , подчёркивающему градиент яркости, модифицированным оператором Робертса [12] (рис. 2, *e*), затем бинаризовать ( $GBW$ ) (рис. 2, *f*) и логически сложить с  $BW$ . В результате будем иметь связное изображение силуэта человека (рис. 2, *g*). Если в  $GBW$  «залить дыры», то получим изображение силуэта (рис. 2, *h*), аналогичное силуэту на рис. 2, *g*, но без теней в области ног. Из приведённого примера следует, что при использовании дополнительного изображения  $GBW$  можно восстановить связность бинарного изображения объекта, а если позволяет текстура, то исключить слабую тень (рис. 2, *h*).

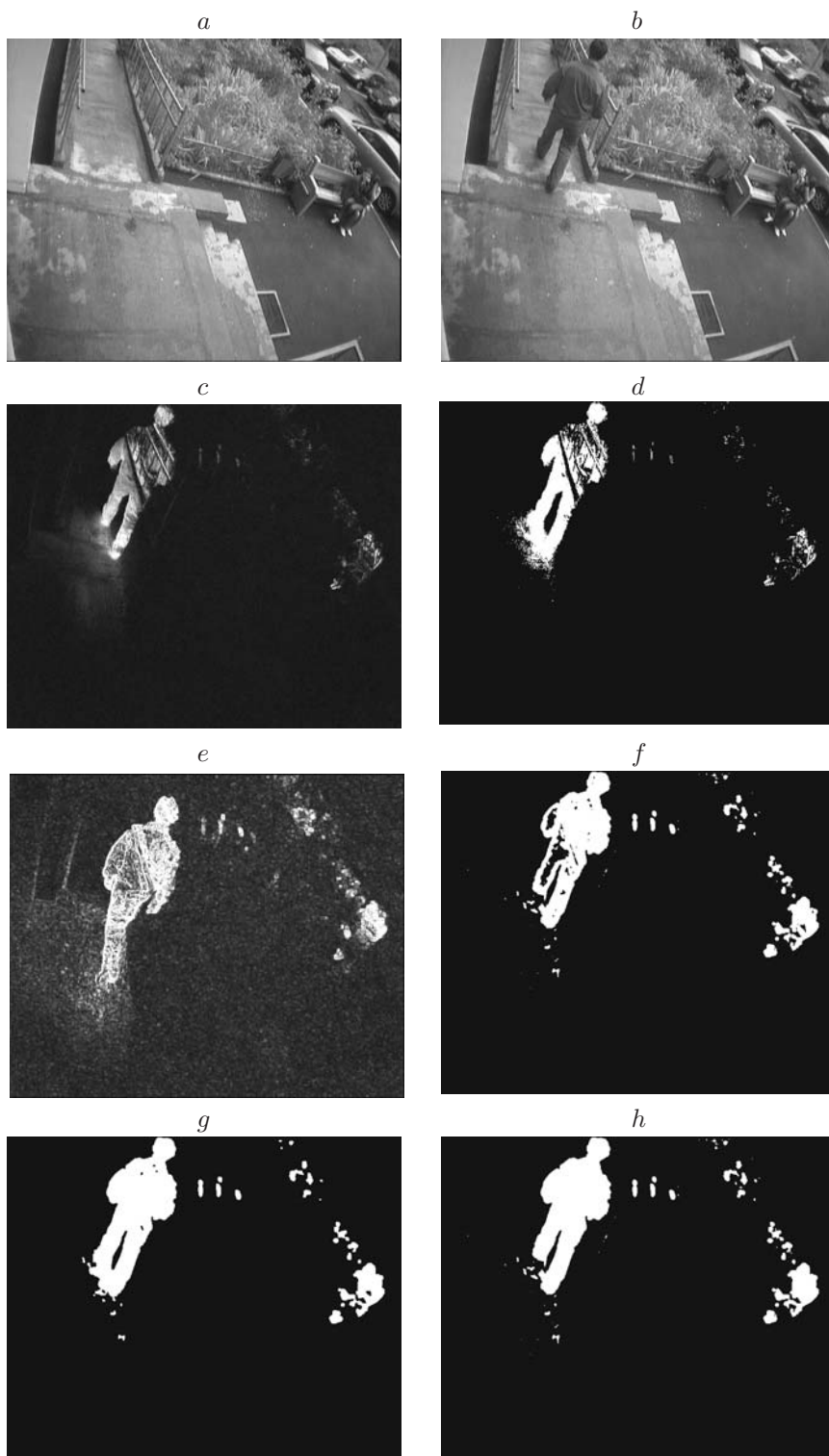
Положение человека относительно ориентации камеры значительно влияет на параметры, вычисляемые по выделенному силуэту. Если человек идёт на камеру или от камеры и падает в направлении движения, то параметры изменяются несущественно. В этом случае по имеющимся параметрам невозможно найти изменения состояния (можно оценить только скорость падения человека). Здесь необходима, как минимум, вторая камера, оптическая ось которой перпендикулярна оси первой камеры. Пример такой ситуации будет приведён в вычислительном эксперименте.

**О параметрах бинарного изображения.** После получения бинарного изображения необходимо подавить шумы и выделить связные объекты. Используя операции математической морфологии, можно убрать объекты, состоящие из нескольких точек (операция эрозии), восстановить площадь (операция дилатации) и заполнить дыры в связных объектах. Такая фильтрация незначительно изменяет площадь и форму объектов, но существенно облегчает дальнейшую обработку. После фильтрации выделяются связные изолированные объекты и определяются их параметры. По силуэтам бинарных объектов могут быть вычислены следующие параметры объекта в текущем кадре  $t$ .

Первичные параметры:

1. Координаты центра тяжести (ЦТ)

$$C_x(t) = \left( \sum_{x, y \in \Omega} x \right) / N, \quad C_y(t) = \left( \sum_{x, y \in \Omega} y \right) / N,$$

*Рис. 2*

где  $N$  — число точек (пикселей), принадлежащих объекту;  $\Omega$  — множество точек объекта, и скорости

$$V_x(t) = C_x(t) - C_x(t-1), \quad V_y(t) = C_y(t) - C_y(t-1).$$

2. Площадь  $A(t) = N$ .

3. Описывающий прямоугольник с параметрами  $bb(t) = [x_{\min}, y_{\min}, W, H]$  и первая разность

$$db(t) = \sqrt{(W(t) - W(t-1))^2 + (H(t) - H(t-1))^2},$$

где  $H, W$  — высота и ширина прямоугольника.

4. Оси инерции  $A_{\min}(t), A_{\max}(t)$ .

5. Угол наклона  $FI(t)$  максимальной оси инерции  $A_{\max}(t)$  относительно оси  $X$  и скорость его изменения

$$VFI(t) = FI(t) - FI(t-1).$$

6. Координаты точек границы  $[G_x, G_y]$ , периметр  $PL(t)$  (длина границы).

7. Проекции объекта  $P_x(t), P_y(t)$  на оси  $X$  и  $Y$ , их вторые моменты  $DX(t), DY(t)$ , которые могут быть получены из  $FI(t)$  и проекций  $A_{\min}(t), A_{\max}(t)$  на оси  $X$  и  $Y$ .

На основе первичных вычисляются параметры, которые в дальнейшем могут использоваться при распознавании ситуаций.

Вторичные параметры:

1. Коэффициент заполнения

$$K_A(t) = A(t)/(W(t)H(t)).$$

2. Коэффициент формы

$$KF(t) = SP(t)/(\pi\sqrt{A(t)}).$$

3. Отношение ширины и высоты описывающего прямоугольника

$$R_{bb}(t) = W(t)/H(t).$$

4. Отношение вторых моментов центроида

$$R_{\text{blob}}(t) = DX(t)/DY(t).$$

5. Отношение осей инерции

$$R_{aa}(t) = A_{\min}(t)/A_{\max}(t), \quad R_{bba}(t) = w_a/h_a,$$

где  $w_a, h_a$  — ширина и высота прямоугольника, вычисленные по  $A_{\min}(t), A_{\max}(t), FI(t)$ .

6. Функция расстояний от ЦТ до точек границы объекта [11]

$$RG = \sqrt{(G_x - C_x)^2 + (G_y - C_y)^2},$$

а также локальные максимумы этой функции, которые можно использовать для распознавания и анализа активности.



Рис. 3

Точечные оценки параметров (к ним относятся  $bb$ ,  $RG$ ,  $[G_x, G_y]$ ,  $db$ ) в вычислительном смысле проще, но менее устойчивы. Остальные параметры носят интегральный характер (они вычисляются по всем точкам объекта) и поэтому более устойчивы. Площадь  $A$  при больших значениях ( $A > 1000$ ) — достаточно устойчивая величина. Периметр  $SP$  имеет значительные ошибки, а коэффициент формы  $KF(t)$  можно использовать в качестве показателя изрезанности границы.

Для видеопоследовательности из 145 кадров (шесть кадров которой показаны на рис. 3) приведены графики параметров  $A$ ,  $C_y$ ,  $R_{bb}$ ,  $R_{aa}$  на рис. 4,  $a$  и  $R_{bba}$ ,  $FI$ ,  $VFI$ ,  $V$  на рис. 4,  $b$ . Из рис. 4,  $a$  видно, что эти параметры неадекватно характеризуют поведение человека, а на рис. 4,  $b$  — они более информативны и позволяют чётко выделить состояния объекта и области перехода из одного состояния в другое.

**Распознавание объектов.** На первом этапе необходимо разделить множество выделенных объектов на два класса (человек и другие объекты). Эту задачу можно решить на основе площади объекта, если известны масштабные коэффициенты, зависящие от положения объекта в кадре. Иначе необходимо использовать набор параметров, который с наибольшей вероятностью проводил бы разделение. Таким набором могут быть параметры  $A$ ,  $R_{bba}$ ,  $R_{aa}$ ,  $FI$ .

Покадровая классификация состояния человека (идёт, стоит, упал, лежит, сидит и пр.) существенно зависит от параметров, которые оцениваются по бинарному изображению. Нестабильность характеристик сцены и объектов обусловлена шумом, освещением, погодными условиями, движением объектов (как объекта в целом, так и его частей) и приводит к широкой вариации оцениваемых параметров и неверному распознаванию при классификации в каждом кадре последовательности. Поэтому в данной работе предлагается на заданном интервале наблюдения выделять интервалы стабильности параметров объекта и оценивать состояние объекта по средним значениям параметров на этих интервалах, что должно повысить вероятность правильной классификации.

**Анализ среднеквадратического отклонения (СКО) на скользящем интервале для выделения интервалов стабильности (постоянства) параметров.** Выде-

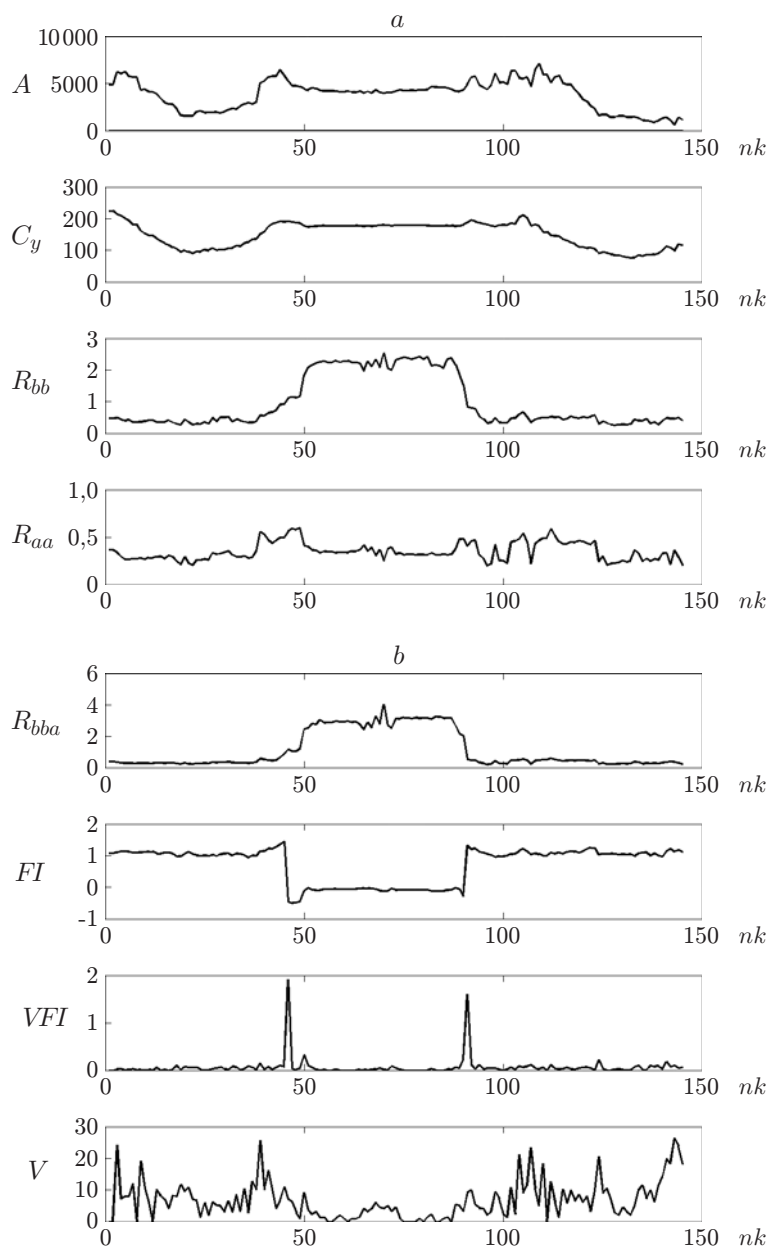


Рис. 4

ление интервалов стабильности параметра необходимо проводить по последовательности значений параметра  $P$ , который более «резко» выделяет такие интервалы. Для  $P$  на заданном интервале  $2L + 1$  вычисляем среднее значение и СКО для всех  $i \in T$  ( $T$  — интервал наблюдения) за исключением  $L$  точек на концах интервала:

$$MP(i) = \sum_{j=i-L}^{j=i+L} P(j)/2L, \quad s_P(i) = \sqrt{\sum_{j=i-L}^{j=i+L} (P(j) - MP(i))^2/2L}, \quad i = \overline{L, T-L},$$

где  $i$  — центр интервала. При наличии резкого изменения  $P$  величина  $s_P(i)$  также резко возрастает и максимум этой величины будет соответствовать точке перехода в другое состояние [8]. Выбрав порог  $P_{rs}$  для  $s_P(i)$ , можно выделить интервалы стабильности па-

Таблица 1

Параметры	Состояния					
	идёт	стоит	лежит	лежит, шевелится	промежуточное (падает, встаёт)	другие
$VC$	$> V_P$	$< V_P$	$< V_P$	$> V_P$	$> V_P$	не определено
$R_{bba}$	$< 1 - \varepsilon$	$< 1 - \varepsilon$	$> 1 + \varepsilon$	$> 1 + \varepsilon$	$\in (1 \pm \varepsilon)$	$\in (1 \pm \varepsilon)$
$FI$	$> 45^\circ + \delta$	$> 45^\circ + \delta$	$< 45^\circ - \delta$	$< 45^\circ - \delta$	$\in (45^\circ \pm \delta)$	$\in (45^\circ \pm \delta)$
$KL$	6	5	4	3	2	1

*Примечание.*  $V_P$  — порог по скорости ЦТ  $VC$ ;  $\varepsilon$  — величина зоны неопределённости по параметру  $R_{bba}$ ;  $\delta$  — величина зоны неопределённости по параметру  $FI$ .

раметров и интервалы изменения состояния объекта.

Скользящее СКО более пригодно для выделения зоны смены состояния (точки смены состояния), так как оно более устойчиво к шумам, хотя ширина пика может быть больше зоны изменения состояния параметра объекта (зависит от величины интервала  $L$ ).

**Классификация состояния.** После выделения интервалов стабильности и промежуточных состояний параметров вычисляем на этих интервалах средние значения, по которым можно разделить поведение объекта на классы в интервале наблюдения.

В качестве параметров были выбраны (как наиболее информативные) средние значения  $VC = \sqrt{(V_x)^2 + (V_y)^2}$ ,  $R_{bba}$ ,  $FI$  на выделенных интервалах стабильности. В табл. 1 приведены значения параметров, соответствующие состояниям объекта, которые можно использовать для классификации, и номера классов состояния объектов ( $KL$ ).

**Результаты вычислительного эксперимента.** Для вычислительного эксперимента взята последовательность из 1217 кадров (шесть её кадров представлены на рис. 3). Обработывался каждый седьмой кадр (всего обработанных кадров 145), в качестве фона

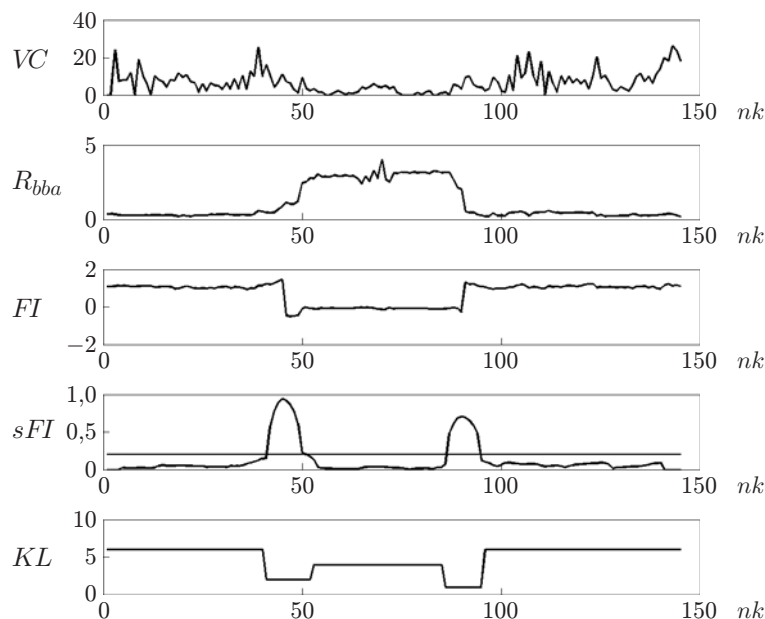


Рис. 5



Таблица 2

Интервал	Параметры			
	$VC$	$R_{bba}$	$FI$	$KL$
1–40	7,8959	0,3071	1,0623	6 (идёт)
41–52	6,7096	1,2504	0,3657	2 (падает)
51–86	1,7836	2,9889	-0,0759	4 (лежит)
85–95	4,8031	1,5165	0,5389	1 (не определено)
94–145	9,2719	0,3507	1,0745	6 (идёт)

использован десятый кадр, не содержащий объектов (начало с 50-го кадра).

На рис. 5 представлены динамика параметров  $VC$ ,  $R_{bba}$ ,  $FI$ ,  $sFI$  и номера классов состояния объекта  $KL$  ( $Pr = 25$ ,  $Prs = 0,2$ ; параметры для классификации:  $V_P = 3$ ,  $\delta = 0,3$ ,  $\varepsilon = 0,15$ ) в зависимости от номера текущего кадра ( $nk$ ). Величина скользящей дисперсии  $sFI$  вычислялась по интервалу  $2L + 1 = 9$  для параметра  $P = FI$ .

В табл. 2 сведены средние значения параметров и номера классов состояния объекта на интервалах стабильности.

На рис. 6 показаны четыре кадра последовательности из 3980 кадров. Обработывался каждый 20-й кадр (начало с 300-го кадра), в качестве фона использовался десятый кадр, не содержащий объектов. Динамика параметров и номера классов состояния объектов на рис. 7 соответствуют рис. 5, кроме величины  $Prs = 0,15$ .

В табл. 3 представлены средние значения параметров и номера классов состояния объекта на интервалах стабильности.

Из результатов второго эксперимента (см. рис. 7 и табл. 3) следует, что при дви-



Рис. 6

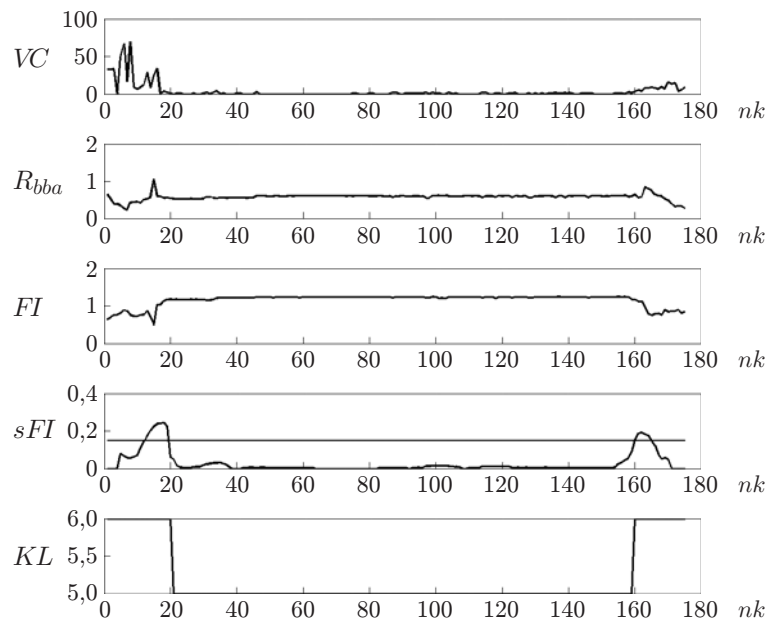


Рис. 7

Таблица 3

Интервал	Параметры			
	$VC$	$R_{bba}$	$FI$	$KL$
1–12	28,3278	0,4159	0,7620	6 (идёт)
13–20	12,8594	0,6086	0,9309	6 (идёт)
21–160	0,5709	0,5893	1,2241	5 (стоит)
161–166	6,4863	0,6842	0,9603	6 (идёт)
167–175	9,1937	0,4509	0,8344	6 (идёт)

жении человека в направлении от камеры (камера одна) и при его падении в том же направлении данный подход не даёт возможности обнаружить упавшего человека — он распознаётся как стоящий. Поэтому необходимо устанавливать камеру так, чтобы ошибки были минимальными. Другой вариант заключается в установке двух камер, что приведёт к усложнению (аппаратному, алгоритмическому и программному) системы наблюдения.

**Заключение.** В предлагаемой работе проведён анализ алгоритмов детектирования упавшего человека по видеопоследовательности, получаемой неподвижной камерой наблюдения. Предложен вариант получения более достоверного силуэта объекта. Проведён анализ параметров выделяемых объектов и выбраны наиболее информативные для распознавания их состояния. Предложена методика повышения устойчивости параметров на основе оценивания их средних значений на выделенных интервалах стабильности, что позволяет повысить достоверность идентификации состояния объектов. Результаты вычислительных экспериментов подтвердили обоснованность предлагаемой методики.

## СПИСОК ЛИТЕРАТУРЫ

1. Luštrek M., Kaluža B. Fall detection and activity recognition with machine learning // Informatica. 2009. **33**, N 2. P. 205–212.

2. **Noury N., Fleury A., Rumeau P. et al.** Fall detection — principles and methods // Proc. of the 29th IEEE Intern. Conf. Engineering in Medicine and Biology Society. Leon, France, 22–26 August, 2007. P. 1663–1666.
3. **Toreyin B. U., Dedeoglu Y., Cetin A. E.** HMM based falling person detection using both audio and video // Proc. of the 14th IEEE Conf. Signal Processing and Communications Applications. Antalya, Turkey, 17–19 April, 2006. P. 1–4.
4. **Elhabian S., El-Sayed K., Ahmed S.** Moving object detection in spatial domain using background removal techniques — State-of-art // Recent Patents on Computer Sci. 2008. **1**. P. 32–54.
5. **Cheung S.-C. S., Kamath C.** Robust background subtraction with foreground validation for urban traffic video // EURASIP Journ. Appl. Sign. Process. 2005. **14**. P. 2330–2340.
6. **He Y., Yang X.** Vehicle detection and tracing based on phase-correlation // Chin. Opt. Lett. 2004. **2**, N 6. P. 323–324.
7. **Нужный С. П., Червяков Н. И.** Детектор движения в цифровой системе охранного видеонаблюдения // GraphiCon'2007. Москва, Россия, 23–27 июня 2007.  
URL: [http://www.graphicon.ru/2007/proceedings/Papers/Paper\\_7.pdf](http://www.graphicon.ru/2007/proceedings/Papers/Paper_7.pdf) (дата обращения: 1.12.2010).
8. **Иванов В. А., Киричук В. С.** Выделение движущихся объектов по последовательности изображений в реальном времени // Автометрия. 2009. **45**, № 5. С. 14–21.
9. **Cucchiara R., Grana C., Piccardi M., Prati A.** Detecting moving objects, ghosts, and shadows in video streams // IEEE Trans. Pattern Anal. Mach. Intell. 2003. **25**, N 10. P. 1337–1342.
10. **Irie K., McKinnon A., Unsworth K., Woodhead I.** Noise-bound method for detecting shadow-free scene changes in image sequences // JOSA A. 2010. **27**, N 2. P. 167–173.
11. **Ekinci M., Gedikli E.** Silhouette-based human motion detection and analysis for real-time automated video surveillance // Turk. Journ. Electr. Eng. and Comp. Sci. 2005. **13**, N 2. P. 199–230.
12. **Kulikov V. A., Popov S. A., Kirichuk V. S.** Registration and analysis of X-ray images received from the system of radiographic control (SRC) "Sibscan" // Proc. of the IASTED Intern. Conf. on Automation, Control and Information Technology (ACIT'2010). Anaheim — Calgary — Zurich: ACTA Press, 2010. Vol. 691. P. 56–61.

*Поступила в редакцию 22 ноября 2010 г.*

---