

УДК 004.932.2

ВЫДЕЛЕНИЕ ДВИЖУЩИХСЯ ОБЪЕКТОВ ПО ПОСЛЕДОВАТЕЛЬНОСТИ ИЗОБРАЖЕНИЙ В РЕАЛЬНОМ ВРЕМЕНИ

В. А. Иванов, В. С. Киричук

*Институт автоматизации и электрометрии СО РАН,
630090, г. Новосибирск, просп. Академика Коптюга, 1
E-mail: vaivanov@iae.nsk.su*

Предложен новый алгоритм оценивания стационарной части фона наблюдаемой сцены и его адаптации к изменениям содержания сцены и условий наблюдения. Алгоритм является оптимальным в статистическом смысле и использует минимальное количество статистических характеристик (скользящее среднее, дисперсию и медиану), что позволяет проводить их рекуррентное вычисление, необходимое при реализации в реальном времени. Обоснованность предложенной методики подтверждена результатами вычислительного эксперимента с реальными видеопоследовательностями.

Ключевые слова: объект, видеопоследовательность, изображение, фон, интервал времени, детектор, траектория движения.

Введение. Выделение движущихся объектов по последовательности изображений является базовой задачей для охранных систем, систем контроля движения транспорта, контроля опасных (несанкционированных действий) ситуаций и т. д. Так, в [1] приведен обзор подходов к оцениванию фона, необходимого при вычислении разности текущего кадра видеопоследовательности и фона для разделения пикселей на два класса: фон и подвижные объекты. В работе [2] представлен обзор устоявшихся схем удаления фона, методов обнаружения подвижных объектов и их распознавания. Конкретным реализациям детекторов движения на основе вычисления фона посвящены работы [3–5].

Как следует из этих работ, к настоящему времени сложилась структура типовых алгоритмов обработки:

- 1) оценивания стационарной части наблюдаемой сцены (заднего плана сцены) с получением ее статистических характеристик;
- 2) получения разностных изображений разнесенных во времени кадров и разностных изображений последовательности с изображением стационарной части сцены;
- 3) выделения областей, принадлежащих объектам, идентификации этих объектов, определения траекторий движения объектов и их классификации;
- 4) адаптации стационарной части фона к изменяющимся условиям регистрации и изменениям содержания сцены;
- 5) регистрации заданных ситуаций с выдачей необходимых сообщений.

Несмотря на множество предложенных решений (в основном эвристических), на сегодняшний день не удалось разработать универсального алгоритма решения задачи оценивания фона.

Целью данной работы является создание алгоритма оценивания стационарной части фона наблюдаемой сцены и его адаптация к изменениям содержания сцены и условиям наблюдения. Предлагаемый алгоритм является оптимальным в статистическом плане при использовании минимального количества статистических характеристик (скользящего среднего, дисперсии и медианы). Использование минимального набора статистик, допускающих рекуррентный расчет, обусловлено тем, что алгоритм должен работать в реальном времени, т. е. не должен иметь высокую вычислительную сложность.

Постановка задачи. Пусть имеется последовательность (полученная стационарной неподвижной камерой) из N изображений $D_l(i, j)$, $l = 1, N$, которая содержит квазистационарный фон $m(i, j)$ (считаем, что на первых N изображениях фон стационарен) и изображения движущихся объектов. Необходимо оценить значение стационарного фона $m(i, j)$.

Рассмотрим поведение сигнала в некоторой произвольной точке изображения $D_l(i, j)$, предполагая, что за время наблюдения $T = N\Delta t$ (Δt — интервал времени между кадрами последовательности) через эту точку может пройти n объектов (далее индексы (i, j) опустим):

$$D_l = m + \sum_{k=1}^n F_k(p_k, q_k) + \xi_l, \quad (1)$$

где m — стационарный фон, подлежащий оцениванию; (p_k, q_k) — интервал присутствия объекта k ; ξ_l — случайный шум; $F_k(p_k, q_k)$ — форма объекта, неизвестная априори.

Относительно формы объектов предполагаем:

- среднее F_k отличается от нуля и равно m_0 ;
- длительность нахождения объекта в точке изображения $|p_k - q_k|$ неизвестна (считаем, что определен диапазон ее изменения);
- дисперсия σ_F^2 неизвестна.

Если среднее значение яркости объектов равно 0, то наличие объекта вносит незначительный вклад в ошибку оценивания m , следовательно, такие объекты не влияют на оценку m . При $m \neq 0$ происходит существенное искажение оценок $\hat{m} = \bar{D}_l$ или $\hat{m} = \text{median}(D_l)$. Наиболее распространенная оценка \hat{m} через медиану эффективна, когда суммарная длительность объектов

$$Q = \sum_{l=1}^n |p_k - q_k| \ll N.$$

Однако когда Q сравнима с N , медиана дает заведомо несостоятельные оценки. Поэтому единственным выходом из данной ситуации является нахождение «зоны действия» объектов и исключение ее из процедуры оценивания m .

Поиск объектов. В силу того что форма объектов неизвестна, т. е. невозможно применить методы согласованной фильтрации, алгоритм поиска объектов будет базироваться на отличии первых двух моментов (среднего и дисперсии) фона и объектов. Рассчитаем на некотором интервале наблюдения $2L$ среднее значение и дисперсию:

$$\hat{M}_i = \sum_{j=i-L}^{j=i+L} D_j / 2L, \quad s_i^2 = \sum_{j=i-L}^{j=i+L} (D_j - \hat{M}_i)^2 / (2L - 1), \quad (2)$$

где i — центр интервала наблюдения.

Если интервал не содержит объектов, т. е. $D_j = m + \xi_j$, то s_i^2 подчиняется χ^2 -распределению при случайном шуме ξ , близком к нормальному:

$$s_i^2 \in \chi_{2L-1}^2 \sigma_\xi^2 / (2L - 1). \quad (3)$$

Если в интервале наблюдения имеется граница объекта, то s_i^2 будет принадлежать нецентральному χ^2 -распределению с параметрами нецентральности $\lambda = m_0^2(L^2 - \varepsilon^2)$, где ε — отклонение границы объекта от центра интервала наблюдения. Максимум λ достигается, когда граница объекта лежит посередине интервала наблюдения. Следовательно, для

определения точек, принадлежащих фону, необходимо проверить гипотезу о принадлежности s_i^2 к центральному χ^2 -распределению. Равномерно наиболее мощный критерий в классе инвариантных принимает гипотезу, если

$$s_i^2 < C_\alpha \sigma_\xi^2, \quad (4)$$

тогда точка i принадлежит фону. Здесь C_α — квантиль распределения, соответствующая заданному уровню ложных тревог α . Таким образом, алгоритм оценивания точек, принадлежащих фону, строится следующим образом: для всех точек, лежащих в интервале $[L, N - L]$, вычисляются дисперсия и среднее значение, и если $s_i^2 < C_\alpha \sigma_\xi^2$, то считается, что точка принадлежит фону.

Однако если объект проходит через анализируемую точку изображения, имеет постоянную яркость и движется достаточно медленно, так что время его прохождения больше чем $2L \times \Delta t$, то скользящая оценка дисперсии также будет подчиняться χ^2 -распределению и для этой точки последовательности выполняется условие (3), и она будет ложно отнесена к фону. На рис. 1, *a* показана динамика сигнала. Точки с 10-й по 80-ю и с 120-й по 150-ю принадлежат объектам постоянной яркости, при этом условие (3) для них выполняется. На рис. 1, *b* приведена оценка скользящей дисперсии.

Кроме того, если начальный участок принадлежит объекту постоянной яркости, то первый локальный экстремум s_i^2 (либо выполнение условий $s_i^2 < C_\alpha \sigma_\xi^2$ и $s_{i+1}^2 \geq C_\alpha \sigma_\xi^2$) указывает на конец объекта (как на рис. 1, *b* в 80-й точке), а если точки принадлежат фону, то первый локальный экстремум указывает на начало объекта.

Поведение s_i^2 позволяет определить:

1) границы объектов (по переходам $s_i^2 < C_\alpha$ и $s_{i+1}^2 \geq C_\alpha$, $s_i^2 > C_\alpha$ и $s_{i+1}^2 \leq C_\alpha$ или по экстремумам s_i^2);

2) принадлежность точек фону или объекту постоянной яркости.

Поэтому для определения точек, принадлежащих фону, необходимы дополнительные требования.

В реальных условиях наблюдения вероятность принадлежности точки объекту значимо меньше, чем фону. Поэтому полагаем, что базовым интервалом, по которому рассчитывается значение фона, будет конечный интервал длиной не меньше заданного L_{\min}

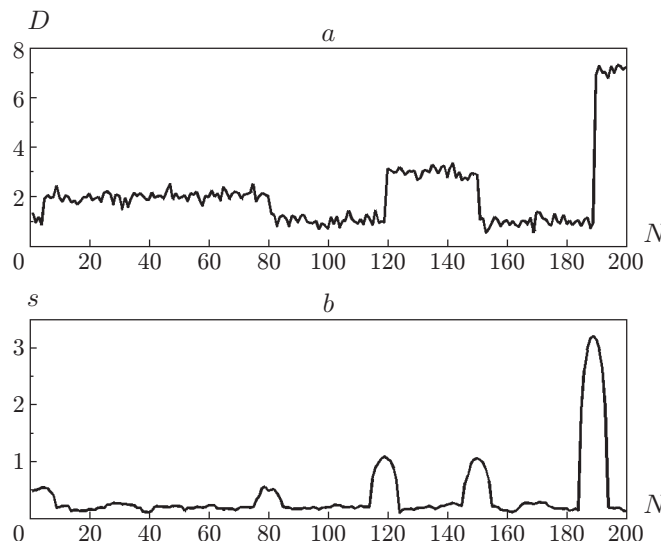


Рис. 1

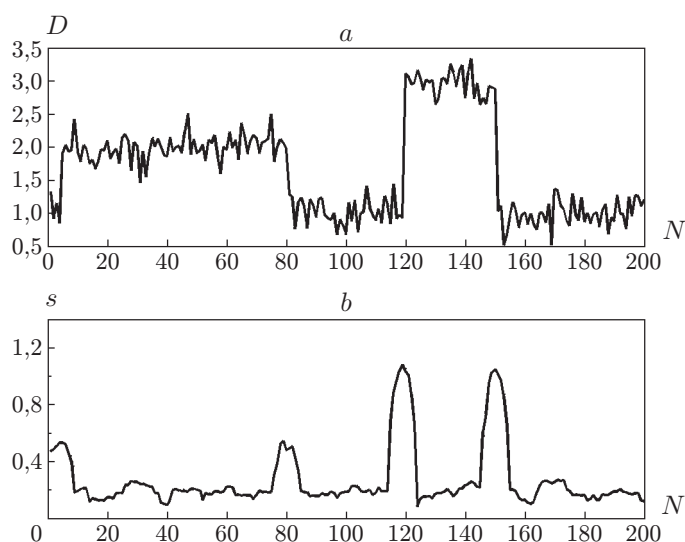


Рис. 2

и выполнение условия $s_i^2 < C_\alpha \sigma_\xi^2$. По найденному базовому интервалу рассчитываем значение стационарного фона как среднее значение или как медиану. При этом существует вероятность того, что этот интервал попадет на объект постоянной яркости и фон будет определен неверно, однако эта ситуация будет исправлена при дальнейшей модификации фона.

При таком определении базового интервала возникают следующие ситуации.

1. Интервал, длительность которого $\geq L_{\min}$, отсутствует в зоне наблюдения. Тогда необходимо продолжить наблюдение до тех пор, пока не появится такой интервал, а если это технически невозможно, то выбираем интервал максимальной длительности из существующих.

2. Выбранный интервал не захватил конечную точку наблюдения. Тогда считаем, что после перехода $s_i^2 < C_\alpha \sigma_\xi^2$ и $s_i^2 \geq C_\alpha \sigma_\xi^2$ (рис. 2) начался объект, и начинаем расчет «стационарности данного объекта».

3. Если $s_i^2 \geq C_\alpha \sigma_\xi^2$ в конце интервала наблюдения, то считаем, что эта точка принадлежит границе (рис. 3), и выбираем интервал $\geq L_{\min}$, ближний к концу или к максимальной длительности из существующих, для расчета значения фона.

Модификация фона и обнаружение оставленных или исчезнувших объектов. Если объект попадает в поле зрения камеры и останавливается на некоторое продолжительное время, то он становится частью фона, и поэтому необходимо менять фон в области нахождения объекта. Аналогичная ситуация возникает и при исчезновении ранее неподвижного объекта из поля зрения камеры. При этом желательно управлять временем, по истечении которого объект считается появившимся или исчезнувшим.

Следовательно, возможны три ситуации:

- точка принадлежит фону;
- точка принадлежит подвижному объекту;
- точка принадлежит медленно движущемуся объекту постоянной яркости.

Для маркировки этих ситуаций для каждой точки поля введем «счетчик стационарности» K_{ij} , который принимает следующие значения:

- 1) $K_{ij} = 0$ — точка принадлежит фону;
- 2) $K_{ij} = -1$ — точка принадлежит подвижному объекту;
- 3) $K_{ij} > 0$ — точка принадлежит объекту постоянной яркости.

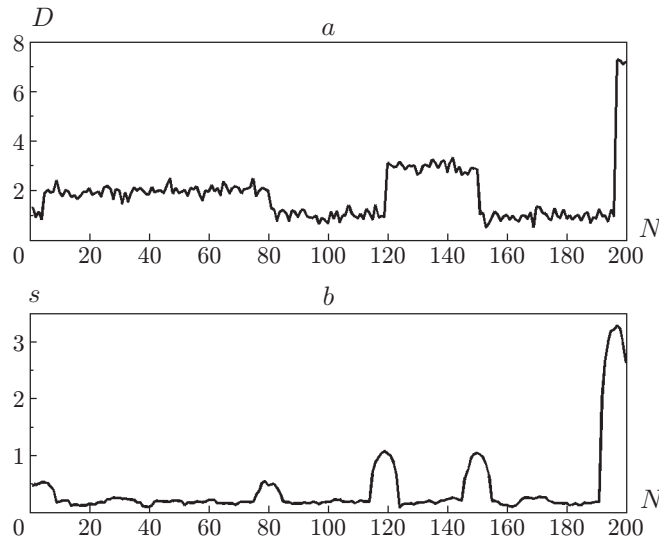


Рис. 3

Динамика изменения K_{ij} определяет поведение фона и выделяет оставленные и исчезнувшие объекты. Начальное значение K_{ij} вычисляется при нахождении базового интервала.

Если базовый интервал захватил конечную точку начального интервала наблюдения N , то $K_{ij} = 0$.

Если базовый интервал не захватил конечную точку (см. рис. 2) и в этой точке $s_{N-L}^2 < C_\alpha \sigma_\xi^2$, то $K_{ij} = 1$ (начался объект), в противном случае $K_{ij} = -1$ (см. рис. 3).

При появлении следующего изображения последовательности (на первом шаге $N + 1$, в общем случае $N + i$) вычисляем s_{N+1-L}^2 и определяем динамику изменения K_{ij} .

Если $s_i^2 \geq C_\alpha \sigma_\xi^2$, то это означает, что интервал длиной $2L$ с центром $N + 1 - L$ захватил границу объекта, поэтому $m_{ij} = m_{ij}$, $K_{ij} = -1$.

Если $s_{N+1-L}^2 < C_\alpha \sigma_\xi^2$, то

1) $K_{ij} = 0$, т. е. данная точка принадлежит фону и необходима его коррекция, например, с помощью фильтра $m_{ij}^{N+1-L} = (1 - \gamma)m_{ij}^{N-L} + D_{N+1-L}(i, j)$, где γ — параметр $\gamma \in (0 - 1)$ [3–5];

2) $K_{ij} = -1$, т. е. если $s_{N+1-L}^2 < C_\alpha \sigma_\xi^2$, то интервал $2L$ имеет постоянную яркость и может быть началом стационарного объекта, а следовательно, $m_{ij} = m_{ij}$, $K_{ij} = 1$;

3) $K_{ij} > 0$ и $K_{ij} < K_0$, т. е. на протяжении K_{ij} кадров точка не меняла свою яркость (здесь K_0 — параметр, определяющий интервал наблюдения, после которого объект считается стационарным), тогда $m_{ij} = m_{ij}$, $K_{ij} = K_{ij} + 1$;

4) $K_{ij} = K_0$, т. е. объект в течение интервала K_0 не менял свою яркость, и считаем, что это стационарный объект, следовательно, $m_{ij} = \bar{D}(i, j)$, $K_{ij} = 0$ (или $m_{ij} = \text{median}(D(i, j))$, $K_{ij} = 0$), где статистика вычисляется по $K_0 + L$ последним изображениям. В таком случае происходит модификация фона, обусловленная появлением или исчезновением объекта.

Апробация данного алгоритма на множестве реальных видеопоследовательностей выявила ситуации, когда алгоритм не может определить стационарные объекты. В случае если объект постоянной яркости медленно удаляется или приближается к камере или стоит на месте, а форма его меняется по формуле (4) (например, человек стоит и машет руками), то центральная часть объекта считается стационарной. Это означает, что применения условия (4) недостаточно и необходимо проверить гипотезу о нестационарности



Рис. 4

объекта, что, в свою очередь, приводит к значительному усложнению алгоритма: приходится переходить от поточечных операций к пространственной двумерной обработке.

Алгоритм проверки объекта на стационарность. Если объект является стационарным, то его форма и параметры не должны меняться в течение всего периода наблюдения K_0 и, кроме того, в силу условия (4) объект на всем участке наблюдения должен иметь хотя бы одну общую точку.

Пусть на T кадре для точки с координатами (l, k) выполнилось условие (4). Тогда, получив разностные изображения

$$\Delta_T(i, j) = D_T(i, j) - m(i, j), \quad \Delta_{T-K_0}(i, j) = D_{T-K_0}(i, j) - m(i, j),$$

осуществим путем линейной, нелинейной и пороговой обработки выделение связных областей Ω_T и Ω_{T-K_0} , содержащих точки (l, k) . Для стационарного объекта эти области должны совпадать: $\Omega_T = \Omega_{T-K_0}$, однако наличие шумов может привести к размытию границ объекта. Поэтому, вводя некоторый параметр стационарности β , считаем, что объект является нестационарным, если зона пересечения множеств Ω_T и Ω_{T-K_0} значимо меньше площади объединения этих множеств, т. е.

$$Q = \frac{\text{Area}\{\Omega_T \cap \Omega_{T-K_0}\}}{\text{Area}\{\Omega_T \cup \Omega_{T-K_0}\}} < 1 - \beta. \quad (5)$$

При идеальных условиях регистрации для стационарного объекта $Q = 1$.

Если гипотеза о нестационарности подтверждается, то $m(l, k) = m(l, k)$, $K(l, k) = -1$ для всех точек, принадлежащих множеству Ω_T . Когда условие (5) не выполняется, объект считается стационарным и при этом $K(l, k) = 0$, т. е. все точки множества Ω_T принадлежат фону и вычисляются по $K_0 + L$ последним изображениям.



Рис. 5

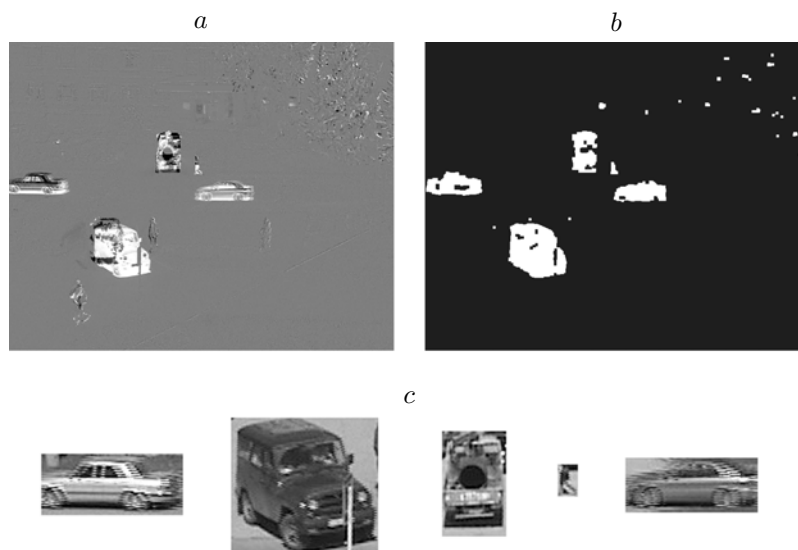


Рис. 6

Результаты вычислительного эксперимента. Для вычислительного эксперимента использовались видеопоследовательности, полученные стационарной камерой SONY DIGITAL VIDEO с разрешением 720×576 , угол зрения 26° , расстояние от камеры до центра сцены ~ 70 м. Проводилось наблюдение участка дороги, где движущимися объектами являлись автомобили и люди. Обработка видеопоследовательности осуществлялась с использованием предложенного алгоритма.

На рис. 4, *a—c* представлены кадры последовательности 1, 50, 100, по которым оценивается начальный фон. На всех кадрах присутствуют движущиеся объекты (автомобили и пешеходы). На рис. 5, *a* приведено изображение стационарного фона, на котором остались только изображения сцены и неподвижных объектов. На рис. 5, *b* показан 150-й кадр, а на рис. 5, *c* — динамический фон после этого кадра. На рис. 6, *a* представлена разность текущего фона и 150-го кадра последовательности, а на рис. 6, *b* — результат порогового выделения движущихся объектов. Алгоритм безошибочно выделил пять разномасштабных движущихся объектов (четыре автомобиля и один пешеход) (рис. 6, *c*). Наличие шумовой компоненты в правом верхнем углу кадра (см. рис. 6, *a, b*) обусловлено колебанием крон деревьев и требует дополнительного исследования для устранения этого явления. Таким образом, результаты вычислительного эксперимента показали обоснованность принятых алгоритмических решений для различных ситуаций детектирования движущихся объектов.

Заключение. В данной работе предложен новый алгоритм оценивания стационарной части наблюдаемой сцены по последовательности видеоизображений. Оригинальность подхода определяется:

- использованием равномерно наиболее мощного критерия в классе инвариантных для проверки гипотез о принадлежности точек изображения объекту или фону;
- введением классификационного признака для областей изображения, определяющего их принадлежность фону, объекту постоянной яркости или объекту переменной яркости;
- введением параметра стационарности объекта, обеспечивающего управляемую модификацию фона.

Вычислительные эксперименты с реальными видеопоследовательностями подтвердили обоснованность предложенной методики.

СПИСОК ЛИТЕРАТУРЫ

1. **Piccardi M.** Background subtraction techniques: A review // <http://www.cs.mcgill.ca/~ethul/pub/course/ecse626/project-report.pdf>
2. **Elhabian S. Y., El-Sayed K. M., Ahmed S. H.** Moving object detection in spatial domain using background removal techniques — state-of-art // Recent Patents on Comput. Sci. 2008. **1**. P. 32–54.
3. **Cheung S.-C. S., Kamath C.** Robust background subtraction with foreground validation for urban traffic video // EURASIP JASP. 2005. N 14. P. 2330–2340.
4. **He Y., Yang X.** Vehicle detection and tracing based on phase-correlation // Chin. Opt. Lett. 2004. **2**, N 6. P. 323–324.
5. **Нужный С. П., Червяков Н. И.** Детектор движения в цифровой системе охранного видеонаблюдения // www.graphicon.ru/2007/proceedings/Papers/Paper_7.pdf

Поступила в редакцию 17 июля 2009 г.
