

# Детектор движения в цифровой системе охранного видеонаблюдения

Нужный С.П., Червяков Н.И.

Кафедра алгебры,

Ставропольский государственный университет, Ставрополь, Россия

{nuzhny007, chervykov}@yandex.ru

## Аннотация

В связи со стремительным распространением цифровых систем видеонаблюдения повышаются требования к качеству обработки и анализа видеоизображений. Сегодня недостаточно констатации движения в кадре и устойчивости детектора к дождю и снегу. Необходимо также выделять и распознавать движущиеся объекты, отслеживать их перемещение, проводить анализ поведения.

В статье предлагается реализация программного детектора движения с возможностью выделения и сопровождения движущихся объектов, а также детектирования оставленных предметов.

**Ключевые слова:** вычитание фона, сегментация, сопровождение объектов

## 1. ВВЕДЕНИЕ

Цифровая система видеонаблюдения на базе персонального компьютера (ПК) представляет собой программно-аппаратный комплекс, состоящий из видеоканалов, устройств захвата видеосигнала и программного модуля. Последний осуществляет обработку, анализ, хранение, отображение видеопотока, поступающего от плат захвата, а также некоторые другие функции.

### 1.1 Система реального времени

Количество видеоканалов в системе видеонаблюдения, а также разрешение захвата и частота кадров захвата на каждом канале могут варьироваться. Максимальная конфигурация системы на одном компьютере ограничивается пропускной способностью PCI-шины и составляет около 8-ми цветных (формат RGB24) видеоканалов с разрешением 384x288 пикселей и частотой 25 кадров/сек. или 3-х каналов с разрешением 768x576 пикселей и частотой 25 кадров/сек. Следовательно, на анализ и сжатие одного кадра с разрешением 384x288 пикселей отводится менее 5 мсек. Построение системы детектирования движения в таких условиях становится достаточно сложной проблемой.

### 1.2 Структура детектора движения

Выделение и сопровождение объектов в видеопотоке осуществляется с помощью моделей разных уровней. Модель более высокого уровня использует данные, полученные от модели предыдущего уровня (рисунок 1). Детектирование оставленных предметов происходит на этапе сопровождения объектов на основании оценки его скорости. В случае положительного детектирования область, занимаемая предметом, становится частью фона. Данный факт отображён на схеме в виде обратной связи.

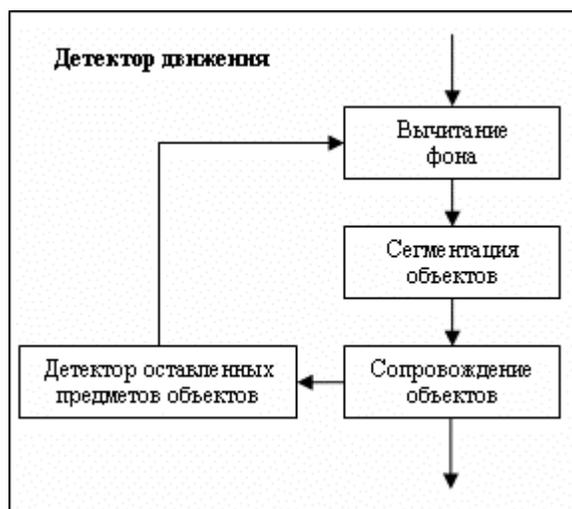


Рис. 1. Структура детектора движения

На **1-м уровне** происходит вычитание фона (background subtraction), то есть классификация пикселей изображения на два класса: пиксели заднего (фон) и переднего плана.

На **2-м уровне** происходит сегментация объектов (objects segmentation) – из множества пикселей переднего плана выделяется множество объектов, то есть семантически связанных областей изображения, принадлежащих одному движущемуся объекту (человеку, автомобилю и др.).

На **3-м уровне** осуществляется сопровождение объектов (objects tracking), то есть межкадровое связывание выделенных областей, вычисление траектории движения, координат, скорости, размера движущихся объектов, их идентификация.

## 2. ВЫЧИТАНИЕ ФОНА

В цифровых системах видеонаблюдения вычитание фона является первичным этапом обработки изображений и служит для классификации пикселей изображения на два класса: пиксели заднего (фон) и переднего плана. Затем на полученных данных строятся такие алгоритмы как выделение, сопровождение и распознавание движущихся объектов.

Разработанный метод вычитания фона является модификацией алгоритмов, описанных в [1] и [4].

Пусть  $C$  – множество пикселей заднего плана,  $D$  – переднего плана. Каждый пиксель заднего плана изображения моделируется с помощью нормально распределённой случайной величины  $\xi$ , то есть

$$C \sim N(\xi, m, \sigma^2) \quad (1)$$

где  $m$  – выборочное среднее значение,  $\sigma$  – среднее квадратичное отклонение.

В зависимости от размерности цветового пространства, в котором представлено изображение, будет меняться и размерность случайной величины. Например, для полутонных изображений  $\xi$  – одномерная случайная величина, для RGB, HSV или YUV пространства – трёхмерная. Для определённости рассмотрим одномерную величину.

#### Алгоритм метода

1. Первые  $n$  кадров происходит обучение модели, то есть оценка параметров  $m$  и  $\sigma$ . Все пиксели классифицируются в задний план. Выборочное среднее и среднее квадратичное отклонение для каждого пикселя вычисляются по формулам:

$$m = \frac{1}{n} \sum_{i=1}^n x_i, \quad \sigma = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - m)^2}, \quad \text{где}$$

$x_i, i = \overline{1, n}$  – значение в пикселе по первым  $n$  кадрам.

2. Для каждого следующего кадра и каждого пикселя:

2.1) Обозначим за  $x$  – значение в текущем кадре пикселя с координатами  $(i, j)$ . Если

$$\frac{|m - x|}{\sigma} \leq e \quad (2)$$

то данный пиксель классифицируется в задний план, иначе – в передний. Параметр  $e$  влияет на чувствительность алгоритма. Его значение получается в

результате решения уравнения:  $P\left(\frac{|m - \xi|}{\sigma} \leq e\right) = a$ ,

где  $a$  – вероятность того, что отклонение, вызванное белым шумом, попадёт в  $e$ -окрестность.

2.2) Если текущий пиксель был классифицирован как пиксель фона, то производим обновление его статистических параметров с помощью низкочастотного фильтра рекурсивного сглаживания:

$$m_t = (1 - \alpha_1)m_{t-1} + \alpha_1 x \quad (3)$$

$$\sigma_t = \sqrt{(1 - \alpha_2)\sigma_{t-1}^2 + \alpha_2(x - m_t)^2} \quad (4)$$

где  $m_{t-1}, \sigma_{t-1}$  – оценки параметров на предыдущем

кадре,  $m_t, \sigma_t$  – текущие оценки,  $\alpha_1, \alpha_2$  – параметры фильтра, отвечающие за скорость обучения модели. Обновление статистики можно производить не с каждым кадром, а через определённые промежутки времени, например, раз в секунду. Частота обновления зависит от наблюдаемой сцены.

2.3) Если текущий пиксель был классифицирован как пиксель переднего плана, то обновление статистики (по формулам 3 и 4) происходит в зависимости от желательной скорости реакции системы на открывшийся фон или объекты, ставшие фоном. Это время составляет обычно 3-5 секунд. Такая задержка также позволяет проводить детектирование оставленных объектов.

3. Классифицированные данные поступают в модели следующих уровней: сегментация, сопровождение, распознавание объектов.

На рисунках 2 и 3 показан результат работы вычитания фона уличной сцены. Съёмка велась цветной аналоговой камерой (400 тв-линий, без внешнего объектива), разрешение кадра 384x288. На дальнем плане работа алгоритма осложняется плывущими облаками и деревьями. Однако, эти области достаточно хорошо вписываются в модель нормального распределения пикселей фона. На рисунке 3 хорошо видны области, соответствующие двум проезжающим автомобилям и идущему по противоположному от наблюдателя тротуару пешеходу.



Рис. 2 Кадр уличной сцены



Рис. 3 Результат действия алгоритма вычитания фона

#### Достоинства алгоритма:

- адаптация к медленному изменению освещённости (смена дня и ночи);
- возможность для детектирования оставленных предметов;
- устойчивость к белому шуму;
- высокая точность определения движущихся объектов;
- детектирование объектов, двигающихся с малой скоростью;
- высокая производительность.

#### Недостатки алгоритма:

- медленная реакция на быстрые изменения освещённости (включение света в помещении, появление солнца из-за тучи и др.);
- нет определения теней;

- нет устойчивости к остальным (кроме белого) шумам изображения, поступающего с устройства видеозахвата.

### 3. СЕГМЕНТАЦИЯ ОБЪЕКТОВ

Сегментация объектов – преобразование множества пикселей переднего плана в множество объектов, то есть семантически связанных областей изображения, принадлежащих одному движущемуся объекту (человеку, автомобилю и др.).

Используется простой и очень быстрый однопроходный алгоритм сегментации, моделирующий каждый наблюдаемый объект в виде прямоугольного региона с границей, имеющей определённую толщину.

На каждом шаге работы алгоритма вычитания фона из пикселей переднего плана формируется множество непересекающихся региона  $R$ , имеющих границу с фиксированной толщиной  $l$ .

1. Если очередной пиксель попадает внутрь или на границу уже существующего региона, то регион расширяется до координат этого пикселя.
2. Если региона, соответствующего пикселю не найдено, то создаётся новый регион размером  $1 \times 1$  с центром в этом пикселе.
3. В случае пересечения 2-х регионов происходит их слияние, то есть создаётся один новый регион, включающий в себя два предыдущих.
4. После обработки всего кадра регионы, размер которых меньше заранее установленного минимума, удаляются.

### 4. СОПРОВОЖДЕНИЕ ОБЪЕКТОВ

Разработан алгоритм сопровождения объектов на основании координат, скорости и времени нахождения объекта на кадре. Входные данные алгоритма – множество регионов, найденных на этапе сегментации объектов.

Каждый объект характеризуется:

- координатами центра  $(x_c, y_c)$ ;
- смещением относительно предыдущего кадра  $(dx, dy)$ ;
- временем жизни  $t_l$  – число кадров, на которых был идентифицирован объект;
- весом  $W$  – актуальность объекта;
- уникальным идентификатором  $uid$ .

#### Алгоритм метода

1. На каждом кадре после работы модели 2-го уровня получен линейный список регионов. Объекты отсортированы по убыванию по времени жизни  $t_l$ . Для каждого объекта выполняется поиск соответствующего ему региона:

1.1) Из всех существующих регионов ищется, чей центр удалён на минимальное расстояние в пикселях от центра объекта. Центр объекта выбирается как  $(x_c + dx, y_c + dy)$ . Расстояние должно быть не больше некоторого заранее установленного минимума

$S_{\min}$ . Обновляется статистика объекта:

а) центр:  $(x_c, y_c) = (x_r, y_r)$ ;

б) вес:

$$w_t = (1 - \alpha) \cdot w_{t-1} + \alpha \quad (5)$$

где  $\alpha$  – некоторый коэффициент, лежащий в интервале  $(0, 1)$ . Чем меньше значение этого коэффициента, тем дольше будет данный объект существовать в модели;  
в) время жизни объекта увеличивается на единицу:  
 $t_l = t_l + 1$ .

Найденный регион удаляется из списка.

1.2) Если для объекта регион не найден, то пересчитываются параметры объекта:

а) вес:

$$w_t = (1 - \alpha) \cdot w_{t-1} \quad (6)$$

Если вес  $W_t < W_{\min}$ , то объект считается устаревшим и удаляется из списка;

б) координаты центра:

$$x_{c,t} = x_{c,t-1} + dx, \quad y_{c,t} = y_{c,t-1} + dy \quad (7)$$

2. Для регионов, которым не соответствует ни один из существующих объектов, создаётся новый объект с центром, совпадающим с центром региона  $(x_c, y_c) = (x_r, y_r)$ ,  $dx = 0$ ,  $dy = 0$ ,  $w = w_{def}$ ,  $t_l = 1$ .

3. При дальнейшей работе модели (обводке, распознавании) объекты с временем жизни меньше некоторого порога  $t_l = t_{\min}$  не используются. Такой подход позволяет отсеивать импульсные шумы, а также снег и дождь. Это модификация алгоритма двухпороговой временной селекции объектов, предложенной в работе [2].



Рис. 4. Каждому отслеживаемому объекту присваивается уникальный номер

### 4.3 Детектор оставленных предметов

Предмет считается оставленным, если на последних  $k$  кадрах координаты его центра не изменились или изменились на некоторую малую величину, то есть:

$$x_{c,t-k} = x_{c,t-k+1} = \dots = x_{c,t},$$

$$y_{c,t-k} = y_{c,t-k+1} = \dots = y_{c,t}$$

$$\text{или } \sum_{i=t-k}^t dx_i < e, \quad \sum_{i=t-k}^t dy_i < e \quad (8)$$

Данный критерий напрямую применять неудобно, так как для каждого объекта необходимо хранить и обновлять массив из  $2k$  чисел.

Для решения этой проблемы необходимо найти функцию, удовлетворяющую следующим условиям для координаты  $X$  объекта (для координаты  $Y$  аналогично):

1) соответствие условию (8):

$$g_t(dx_{t-k}, dx_{t-k+1}, \dots, dx_t) < e, \sum_{i=t-k}^t dx_i < e \quad (9)$$

2) возможность получить значение функции на следующем кадре  $t+1$ , зная лишь значение функции на предыдущем кадре  $t$  и смещение объекта  $dx_{t+1}$ :

$$g_{t+1} = f(g_t, dx_{t+1}) \quad (10)$$

Приближением такой функции  $g$  может служить:

$$g_0^* = 0, \quad g_{t+1}^* = \beta \cdot (g_t^* + dx_{t+1}) \quad (11)$$

где  $\beta \in (0; 1)$  - некоторая константа. Подбирая значения  $e$  и  $\beta$ , устанавливается число кадров  $k$ , которые будут существенно влиять на выход функции:  $k = dx_{\max} \cdot \log_{\beta} e$ .

## 5. ЗАКЛЮЧЕНИЕ

Предложенный алгоритм сопровождения объектов позволяет восстанавливать связь с перемещающимся объектом в том случае, если он на короткое время скрылся из поля зрения камеры.

Тестирование детектора проводилось на компьютерах различной конфигурации с процессорами Intel Celeron 2.0 ГГц и Intel Pentium 3.0 ГГц. **Время работы** детектора движения (вычитание фона, сегментация и сопровождение) на цветном кадре с разрешением 384x288 составляет в среднем 10-12 мсек. Данная величина является достаточно хорошим показателем, однако она не удовлетворяет требованиям, указанным в пункте 1.1.

Планируется дополнить детектор движения модулем распознавания объектов (людей и автотранспорта). Также возможно сопряжение стационарной и поворотной видеокамер для решения задачи автоматического наведения поворотной камеры на перемещающиеся объекты.

## 6. БИБЛИОГРАФИЯ

[1] Гаганов В., Конушин А. Сегментация движущихся объектов в видео потоке, <http://cgm.graphicon.ru>, 2004.

[2] Пешков Н. Н. Математическое моделирование и разработка алгоритмов обнаружения и измерения параметров сторонних объектов в системах наблюдения. Кандидатская диссертация. Ставропольский государственный университет, Ставрополь, 2004.

[3] Д. Форсайт, Ж. Понс Компьютерное зрение. Современный подход. : Пер. с англ. – М.: Издательский дом «Вильямс», 2004.

[4] Chris Stauffer, W. Eric L. Grimson Learning Patterns of Activity Using Real-Time Tracking, 2000.

## Об авторах

Нужный Сергей Петрович аспирант кафедры алгебры физико-математического факультета Ставропольского государственного университета. К основным научным интересам относится машинное зрение, классификация, обработка изображений.

Адрес: Ставрополь, 355009, ул. Пушкина, д. 1, СГУ, факультет ФМФ, кафедра алгебры.

Контактный телефон: +79283156442, e-mail: [nuzhny007@yandex.ru](mailto:nuzhny007@yandex.ru)

Червяков Николай Иванович доктор технических наук, профессор, заведующий кафедры прикладной математики и информатики Ставропольского государственного университета. Научные интересы: цифровая обработка сигналов, изображений, распознавание образов, теория чисел, модулярная арифметика, нейроматематика, параллельные вычислительные системы, нейрокомпьютеры.

E-mail: [chervykov@yandex.ru](mailto:chervykov@yandex.ru)

## The motion detector in digital system of security video observation

### Abstract

The quality of processing and the analysis of the video is raising because of prompt distribution of digital systems of video observation requirements. Today there is not enough ascertaining movement in the staff and stability of the detector to a rain and a snow. It is necessary to allocate and recognize also moving objects, to trace their moving, to analyse the behavior.

In this paper is offered the realization of the program motion detector with an opportunity of allocation and support of moving objects, and also detecting the left subjects.

**Keywords:** background subtraction, segmentation, objects tracking.

### About the author(s)

Nuzhny Sergey is an aspirant of Stavropol State University, Department of Physics and Mathematics. His contact email is [nuzhny007@yandex.ru](mailto:nuzhny007@yandex.ru).

Main scientific interests lie in the fields of computer vision, classification and image processing.

Chervykov Nikolay Ivanovich is an professor of Stavropol State University, Department of Physics and Mathematics. His contact email is [chervykov@yandex.ru](mailto:chervykov@yandex.ru).

Main scientific interests lie in the fields of signals and image processing, objects recognition, theory of number, modular arithmetic, neurocomputers.