

Проблемы выделения объектов в компрессированном потоке изображений

П. В. Харебов, С. В. Новиков

Физический факультет ПГУ

Пермский Государственный Университет, Пермь, Россия

kharebov.p@yandex.ru, novikov@networkvideo.ru

Аннотация

Обнаружение и отслеживание объектов требует устойчивой работы алгоритмов моделирования фона. Устойчивость необходима как к шуму, так и к изменениям освещения. Кроме того, стандарты компрессии, используемые в системах охранного телевидения, вносят дополнительные артефакты, затрудняющие анализ видеопотока.

Для отслеживания объектов используется техника «вычитания фона». Для автоматического выбора порога, устойчивого к артефактам, вносимым сжатием, используется метод блочного анализа шума.

Ключевые слова: Видеонаблюдение, вычитание фона, автоматический выбор порога.

1. ВВЕДЕНИЕ

Компрессия видеопотоков в системах охранного телевидения неизбежна из-за необходимости хранения и передачи большого объема данных. Так, например, IP камеры компрессируют видеопоток внутренними средствами перед отправкой по сети.

Существуют две категории стандартов сжатия: покadresные методы (MJPEG) и межкадровые, основанные на оценке движения (MPEG-2, MPEG-4, H.264). Каждый из алгоритмов вносит характерные артефакты в изображение.

Стандарт MJPEG сжимает каждый кадр при помощи алгоритма JPEG. Сжатие приводит к потере мелких деталей изображения и сглаживанию шума.

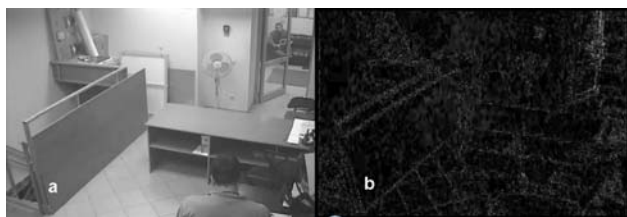


Рис. 1: Разность между последовательными кадрами.

Например, на рис. 1(a) показан JPEG-сжатый кадр, полученный с камеры охранного телевидения. На рис. 1(b) – разность между двумя последовательными JPEG-сжатыми кадрами. Рис. 1(b) дает характерное представление о шумах при JPEG сжатии – отклонение значений пикселей на границах больше чем на однородных областях.

Для видео, сжатого с использованием компенсации движения, кроме описанных выше шумов на краях изображений, появляющихся при JPEG-сжатии, добавляется еще один характерный артефакт – «пульсирующий шум», вызванный

сглаживанием движения в пределах группы кадров. Разница максимальна между последним разностным кадром в группе и новым опорным кадром.

Решаемая нами задача – выделение объектов и отслеживание их перемещений в компрессированном видеопотоке. Пусть дан поток изображений $I^t(i, j)$, где $t = 0, 1, 2, \dots$ – порядковый номер, i, j – координаты пиксела в данном изображении. Элемент изображения (пиксел) – задается значением яркости (уровень серого) в данной точке.

Техника «вычитания фона» широко распространена для выделения объектов в видео [1]. При этом каждый пиксел изображения классифицируется как принадлежащий переднему или заднему плану путем сравнения текущего изображения $I^t(i, j)$ с фоном $B^t(i, j)$. Главным критерием выбора метода «вычитания» была скорость вычисления, что критично для систем видеонаблюдения, работающих в реальном времени и обрабатывающих несколько потоков одновременно. Как показано в [2], наиболее быстрым является метод, при котором фон $B^t(i, j)$ представляет собой усредненное по времени значение $I^t(i, j)$. Пиксель (i, j) считается принадлежащим переднему плану, если $|I^t(i, j) - B^t(i, j)| > \tau$, где τ – некоторый порог.

Возникает проблема выбора порога сравнения. Использование глобального для изображения порога, не позволяют артефакты, вносимые сжатием – порог сравнения τ должен быть своим у отдельного пиксела. Кроме того, шумы не постоянны, что приводит к необходимости изменения порога τ не только в пространстве, но и во времени.

Цель данной работы – разработка метода автоматического выбора порога τ для каждого пиксела, учитывающий историю шума.

2. АНАЛИЗ ШУМА

Шумовые отклонения текущего значения пиксела от фонового значения, присутствуют всегда. Для поиска амплитуды шума изображение разбивается на блоки размером $P * P$ пикселей. В каждом блоке ищется максимальное отклонение между двумя последовательными кадрами:

$$D^t(k, l) = \max_{i, j} (|I^t(i, j) - I^{t-1}(i, j)|), \quad i = \overline{k, k+P}, \quad j = \overline{l, l+P}$$

В течение N кадров ищется максимальное значение пиксела блока:

$$M^t(k, l) = \max_v (D^v(k, l)), \quad v = \overline{t, t-N+1},$$

причем, для учета «пульсирующего шума», N должно быть не меньше числа кадров в группе $N_1 : N \geq N_1$.

Т.к. отклонение может быть вызвано движением, слишком большие значения $M^f(k,l)$ обрезаются так, что если

$$M^f(k,l) > \frac{3}{2}H^{f-1}(k,l), \text{ то } M^f(k,l) = \frac{3}{2}H^{f-1}(k,l).$$

Обновляется матрица истории шума $H^f(k,l)$:

$$H^f(k,l) = \beta M^f(k,l) + (1 - \beta)H^{f-1}(k,l).$$

Где β – параметр, отвечающий за скорость обновления $H^f(k,l)$. $H^f(k,l)$ характеризует возможное отклонение пиксела от фоновое значения.



Рис. 2: История шума $H^f(k,l)$.

На рис. 2 показана история шума при $P = 8$. Для удобства восприятия значение каждого пиксела умножено на 20. Кадр соответствует рис. 1 (а). Видна неравномерность шума значений пикселей. Использование глобального фиксированного порога, при сравнении текущего кадра с фоном, привело бы к ложным срабатываниям в блоках с большим шумом и пропущенными срабатываниями в не шумящих блоках.

3. ВЫДЕЛЕНИЕ ПЕРЕДНЕГО ПЛАНА

Классификация пикселей на передний и задний план обозначается с помощью маски пикселей переднего плана $F^f(i,j)$:

$$F^f(i,j) = \begin{cases} 1, & \text{если } |I^f(i,j) - B^{f-1}(i,j)| > aH\left(\frac{i}{P}, \frac{j}{P}\right) + b \\ 0, & \text{иначе} \end{cases}$$

Где параметры a , b выбираются таким образом, чтобы при заданных N и P обеспечить необходимый уровень верных срабатываний.

$F^f(i,j)$ сглаживается с помощью операций математической морфологии.

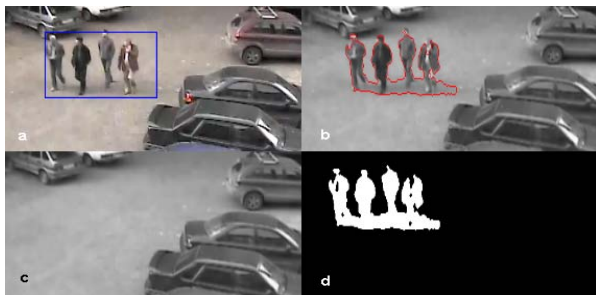


Рис. 3: Выделение переднего плана.

На рис. 3(а) показано исходное изображение и выделенный объект, рис. 1(б) – анализируемое изображение $I^f(i,j)$, рис. 1(с) – фон $B^f(i,j)$, рис. 1(д) – маска пикселей переднего плана $F^f(i,j)$.

Последний шаг – обновление фона в пикселях заднего плана:

$$B^f(i,j) = \alpha B^{f-1}(i,j) + (1 - \alpha)I^f(i,j)$$

4. ЗАКЛЮЧЕНИЕ

Предложенный метод устойчиво работает не только в пространстве яркости, но и с обработанными градиентными методами изображениями, что позволяет отслеживать объекты и в условиях сильно флукутирующего освещения.

Реализована на практике система выделения и отслеживания объектов в компрессированном потоке изображений. Система работает в реальном времени (25 к/с) для восьми камер с разрешением 640x480 на компьютере с процессором Intel Pentium Core 2 Quad.

Достигнут компромисс между устойчивостью вычислительно сложных методов вычитания фона (например, моделирование с помощью гауссовских распределений[1,2]) и скоростью выполнения простейших методов.

5. СПИСОК ЛИТЕРАТУРЫ

- [1] R.J. Radke, A. Srinivas, O. Al-Kofahi, and B. Roysam. Image change detection algorithms: a systematic survey. IEEE Trans. Image Processing, vol. 14, no. 3, pp. 294–307, Mar. 2005.
- [2] M. Piccardi. Background Subtraction Techniques: A Review. IEEE SMC/ICSMC, vol. 4, pp. 3099–3104, 2004.

Об авторах

Харевов Петр – аспирант кафедры компьютерных систем и коммуникаций ПГУ. E-mail: kharebov.p@yandex.ru

Новиков Сергей – к.т.н., руководитель проекта «Domination». E-mail: novikov@networkvideo.ru

The problems of object tracking in compressed video stream.

Objects detection and tracking demands a steady work background modeling algorithms. Robustness is necessary both for noise, and for illumination changes. Besides, the compression standards used in surveillance systems create the additional artifacts complicating the video stream analysis.

In this paper the method of the noise block analysis is presented. The technics of background subtraction is used for objects tracking.

Keywords: video surveillance, background subtraction, automatic thresholding.

About the authors

Petr Kharebov – PhD student of computer systems and telecommunications department of Perm State University. E-mail: kharebov.p@yandex.ru.

Sergey Novikov – PhD, the «Domination» project head. E-mail: novikov@networkvideo.ru.