

УДК 681.518.3

Исследование методов и алгоритмов обнаружения движущихся объектов в видеопотоке

*Сюй Лэй, магистрант,
Россия, 105005, г. Москва, МГТУ им. Н.Э. Баумана,
кафедра «Системы автоматического управления»*

*Научный руководитель: Гаврилов А.И., к.т.н., доцент
Россия, 105005, г. Москва, МГТУ им. Н.Э. Баумана,
кафедра «Системы автоматического управления»
pupkov@iu1.bmstu.ru*

Введение

С развитием технологий и повышением общесоциального уровня постоянно повышаются требования к безопасности и осведомленности групп и отдельных лиц, поэтому задача мониторинга нашла свое применение в различных сферах деятельности человека. Существенную роль в решении задачи мониторинга играют системы видеонаблюдения. В настоящее время они широко используются в банках, музеях, на дорогах, в системах коммерческой, военной, общественной безопасности, электро-обеспечении и других областях. Помимо типовых функций - вывода видеоизображения на монитор, многопоточного наблюдения и записи на жесткий диск – в настоящее время системы мониторинга дополняются модулями интеллектуального обнаружения движения и сопровождения цели [1].

В системе обнаружения и мониторинга движущихся объектов немаловажной составляющей является уровень интеллектуальности. Целью данной работы является создание интеллектуального алгоритма, с помощью которого можно исключить фоновую составляющую изображения и точно определить относительные координаты движущейся цели.

На данный момент уже разработаны алгоритмы и системы, позволяющие успешно

решать эту задачу. В данной работе рассмотрены общие процедуры решения задачи обнаружения подвижных объектов, с акцентом на моделирование фона; решена задача программной реализации этих подходов средствами Матлаб. Также в рамках данной работы проведены эксперименты с различными видеопотоками и представлено сравнение этих методов.

Типичными подходами к обнаружению движения являются метод выделения фона, метод межкадровой разности и метод оптического потока.

Межкадровая разность

Вычисление межкадровой разности является распространённым методом первичного обнаружения движения, после выполнения которого можно определить, присутствует ли в потоке кадров движение.

Алгоритм вычисления межкадровой разности выглядит следующим образом: на вход алгоритма поступают два видеокadra, представляющие собой две последовательности байт[2]. Производится вычисление попиксельных межкадровых разностей по следующей схеме

$$d_t(x, y) = I_t(x, y) - I_{t-1}(x, y)$$

Разность сравнивается с заданным порогом. В результате сравнения формируется двоичная маска:

$$m_t(x, y) = \begin{cases} 0, & d_t(x, y) < T \\ 1, & d_t(x, y) \geq T \end{cases}$$

где $m_t(x, y)$ - значение t -го элемента маски, T – порог сравнения, иногда называемый также порогом или уровнем чувствительности.

Метод вычитания фона

Существует много алгоритмов метода вычитания фона, но, в принципе, почти все состоят из процедур, которые представлены на Рис.1. Самыми главными являются предобработка, моделирование фона, обнаружение движения и постобработка.

Типовая схема реализации метода вычитания фона представлена на Рис.1



Рис. 1. Схема обнаружения подвижных объектов

1. Предобработка

Основной задачей [3] на этапе предобработки растровых данных является подготовка изображений к обнаружению на них движущихся объектов. Поэтому предобработка должна сводиться к выделению на изображении тех его элементов, которые предположительно принадлежат движущимся объектам. Пока нам достаточно считать движущимся объектом то, что человек воспринимает, как движущийся объект, наблюдая глазом один кадр видеопотока за другим. Далее мы дадим более точную формулировку движущегося объекта с точки зрения детектора движения.

1) Уменьшение размера кадра

Самым первым шагом на этапе предобработки растровых данных часто является уменьшение размера кадра. Выполнение данной операции на самом начальном этапе даёт два важных преимущества:

а) Существенное сокращение объема информации, которую необходимо обрабатывать на каждом из последующих этапов.

б) Ослабление аддитивного и импульсного шумов на входных данных.

Иначе этот шаг можно назвать переходом к усреднённым значениям пикселей, поскольку значения яркости исходного пикселя заменяются усреднёнными значениями яркость данного пикселя и его соседей.[1] Результирующий пиксель будем называть минимальной зоной (минзоной) исходного изображения. Для случая усреднения четырёх пикселей в одну минзону преобразование запишется так:

$$I'_t(x, y) = \frac{I_t(2x, 2y) + I_t(2x, 2y + 1) + I_t(2x + 1, 2y) + I_t(2x + 1, 2y + 1)}{4}$$

где $I_t(x, y)$ - яркость минзоны из x -ой строки и y -го столбца, другие - значения яркости пикселя.

2) Преобразование формата изображения

Яркость есть характеристика пикселя, которая чаще всего используется при вычитании фона.[2] Поэтому необходимо переводить видеоряд в полутоновое изображение формат, в котором включается яркость. На пример, формат YUV, в которой цвет представляется как 3 компоненты - яркость (Y) и две цветоразностных (U и V)

В системе Матлаб можно конвертировать видеопоток в последовательность полутоновых изображений или формат YCbCr, где Y - компонента яркости, CB и CR являются синей и красной цветоразностными компонентами.

2. Методы моделирования фона

1) Фиксированный фон

а) Первый кадр как фон

Этот метод наиболее прост. Алгоритм сохраняет первый кадр видеопоследовательности, а потом для каждого следующего кадра применяет порог к модулю разности текущего и сохраненного изображения по каждому пикселю.

б) Усреднение

Первый метод редко используется, из-за строгого требования к однородности фона, которое трудно выполняется. Например, на шоссе или магистрали редко бывают ситуации, когда никаких подвижных объектов нет. В этом случае, можно моделировать фон с небольшим [4] количеством (30-100) последовательных кадров видеопотока. Одним из самых простых и эффективных является метод усреднения. Моделью фона будет среднее от первых n кадров:

$$BG(x, y) = \frac{1}{n} \cdot \sum_{i=1}^n I_i(x, y)$$

2) Обновление фона

Для того, чтобы алгоритм был устойчив к изменению фона (например, освещения), необходимо обновлять фон в течении времени на основе начальной модели фона.

а) Приближённый медианный метод

Этот метод очень похож на метод среднего значения. Для подсчета модели фона используется формула:

$$BG_t(x, y) = \begin{cases} BG_{t-1}(x, y) + \delta, & \text{если } I_t(x, y) < BG_{t-1}(x, y) \\ BG_{t-1}(x, y) - \delta, & \text{если } I_t(x, y) > BG_{t-1}(x, y) \\ BG_{t-1}(x, y), & \text{иначе} \end{cases}$$

Здесь δ -заданное значение, которое не должно быть очень маленьким.

б) Метод рекурсивного сглаживания.

Этот метод предполагает плавное изменение модели фона с каждым кадром путем применения низкочастотного фильтра рекурсивного сглаживания. Реализация метода предполагает «смешивание» текущего кадра с моделью фона с коэффициентом α (типично $\alpha = 0.05$ [5]). Затем к разности фона и кадра применяется пороговый фильтр. Можно представить как следующий:

$$BG_t(x, y) = (1 - \alpha)BG_{t-1}(x, y) + \alpha I_{t-1}(x, y),$$

Метод очень прост в реализации, и имеет высокую производительность, применим в условиях плавного изменения освещения, однако он не решает проблему динамического заднего плана.

в) Модификация методов а и б

При применении методов а и б возможно возникать проблема с следом из-за смешивания движущихся объектов на фон при обновлении фона [6]. Чтобы избежать от этого, можно просто обновлять то часть, в которой нет цели, а оставить остальную не измененной. Используется формула:

$$BG_t(x, y) = \begin{cases} \text{обновлять методом а или б,} & \text{если } |BG_t(x, y) - I_t(x, y)| < \tau \\ BG_{t-1}(x, y), & \text{иначе} \end{cases}$$

где τ — порог бинаризации

Таким образом, можно удалить объекты переднего плана в модель фона.

Программная реализация и результаты моделирования

С целью исследования предложенных алгоритмов выполнена программная реализация системы видеомониторинга средствами системы Матлаб 2009а. На компьютере с частотой 2.2 ГГц были проведены опыты с видеорядом с разрешением кадров 240×320 .

На рис.2 представлены результаты моделирования системы.

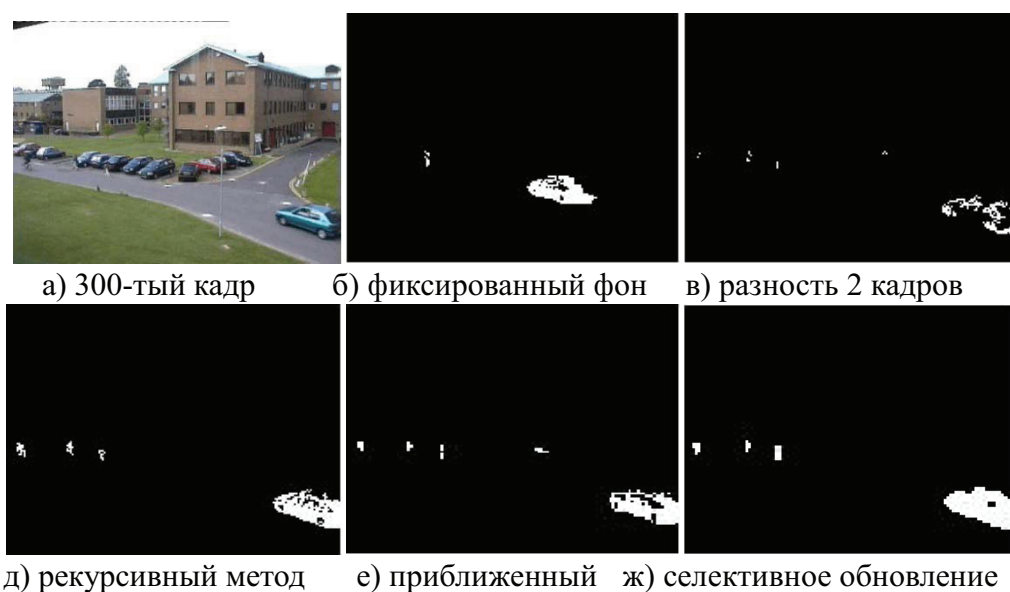


Рис. 2. Результаты моделирования

Результаты сравнения приведены в таблице 1.

Таблица 1

Сравнение методов выделения фона

	Быстродействие(fps)	Точность обнаружения	Нечувствительность к изменению
Фиксированный фон	30	3	1
Межкадровый разность	30	1	5
Приближенный медианная фильтрация	25	4	4
Скользящий среднее	25	4	5
Селективный	12	5	4-5

В результате анализа результатов экспериментов можно сделать следующие выводы :

1) Быстродействие: наилучшим быстродействием обладают методы фиксированного фона и межкадровой разности. Приближенная медианная фильтрация и метод скользящего среднего тоже быстрые, метод гауссовых распределений и селективное обновление - самые медленные.

2) Точность: метод селективного обновления даёт лучший результат, а наихудшими методами является методы гауссового распределений и межкадровой разности.

3) Чувствительность к возмущениям. Оказалось, что все алгоритмы чувствительны к шуму и изменению освещению, но метод межкадровой разности и метод рекурсивного сглаживания наиболее устойчивы к ним.

4) По интегральному критерию, включающему все три показателя эффективности, наилучшими являются приближенный медианный метод

Заключение

В работе были исследованы процедуры обнаружения движущихся объектов методами вычитания фона и определения межкадровой разности.

Разработано алгоритмическое и программное обеспечение и продемонстрировавшее эффективную работу в условия обработки реальных видеопоследовательностей.

На основе предложенных алгоритмов и программного обеспечения можно создать многофункциональную охранную систему, способную обнаруживать объекты на последовательности кадров (в видеопотоке) и отслеживать выделенные объекты как внутри помещения, так и на открытом пространстве.

Отмечено, что почти во всех изображениях присутствуют тени объектов, поэтому в дальнейшей работе будет рассмотрена задача удаления тени. Кроме того, существуют возможности по модификации алгоритмов постобработки с целью улучшения формы движущихся объектов.

Список литературы

1. Yongming Li. An Integrated Traffic and Pedestrian Vision System [Электронный ресурс]. – 1998. URL.<http://www.scs.leeds.ac.uk/imv>
2. Tudan Li. CVOnline: Motion and time sequence Analysis [Электронный ресурс]. – 2002. URL.<http://homepages.inf.ed.ac.uk/rbf/CVonline/motion.htm>
3. Davies D., Palmer P., Mirmehdi M. Detection and Tracking of Very Small Low Contrast Objects [Электронный ресурс]. – 1994. URL. <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.16.196>
4. Вежнев В. Введение в ComputerVision [Электронный ресурс]. – 2003. URL. <http://cgm.computergraphics.ru/content/view/20>
5. Young S., Forshaw M., Hodgetts M. Image comparison methods for perimeter surveillance. - <http://sntbul.bmstu.ru/doc/568760.html>

- University College London Image Processing Group [Электронный ресурс]. – 1999. URL. <http://ieeexplore.ieee.org/stamp/>
6. R.Brooks, C.Breazeal, M.Marjanovic, B.Scassellati. Motion Detection and Segmentation. [Электронный ресурс]. – 1998. URL. <http://repository.cmu.edu/cgi/>