

МЕТОДЫ ВЫЧИТАНИЯ ФОНА ДЛЯ ОБНАРУЖЕНИЯ ОБЪЕКТА В ВИДЕОПОТОКЕ

А.И. Баймухаметова, магистрант

Казанский национальный исследовательский технический университет им.
А.Н. Туполева – КАИ
(Россия, г. Казань)

DOI:10.24412/2500-1000-2022-7-1-74-78

Аннотация. В статье рассматриваются популярные методы вычитания фона для обнаружения объекта в видеопотоке. Вычитание фона – одна из самых важных частей в процессе обработки видеоизображения. При обработке видео необходимо выделять только те объекты, которые имеют определенную ценность и удалить ненужные объекты. Удаление обусловлено тем, что обработка видео с лишними предметами приводит к увеличению времени обработки и требуемой памяти. В этой статье представлено сравнительное исследование нескольких существующих методов вычитания фона, от простого вычитания фона до более сложных статистических методов. Целью данного исследования является поиск сильных и слабых сторон методов, которые широко используются в современном мире.

Ключевые слова: вычитание фона, обработка видео, компьютерное зрение, обнаружение объекта.

Вычитание фона – распространенный метод обработки видео в компьютерном зрении. Областью интереса (ROI – region of interest) в видео может быть несколько объектов: все транспортные средства и люди. Основная цель вычитания фона – это получение последовательных кадров с камеры, обнаружение всех объектов переднего плана как разницы между текущим кадром и изображением статичного фона.

Существует множество методов вычитания фона [1] с различными сильными и слабыми сторонами с точки зрения производительности и вычислительных требований. Надежный алгоритм вычитания фона должен справляться с изменениями освещения, повторяющимися движениями и изменениями погодных условий [2].

Самый простой способ создать фон – это использовать одно изображение в качестве фона в градациях серого/цветного, не содержащее движущихся объектов. Этот кадр может быть сделан в отсутствие движения, или сделан с помощью временного медианного фильтра [3].

Обычные методы.

Обычные методы включают в себя несколько основных подходов: статическая

разница кадров, разница кадров, взвешенное скользящее среднее, взвешенная скользящая дисперсия, адаптивное фоновое обучение, временное среднее, адаптивная медиана Макфарлейна и Шофилда [4], временная медиана Куккиара и др. [3] и Кальдерара и др. [5]. Эти методы могут быть рекурсивными или не рекурсивными подходами.

1) Медианная фильтрация: один из более используемых алгоритмов способа вычитания фона. Он основан на оценке модели фона с помощью вычисления медианного значения каждого входного пикселя. Объект не считается фоном только после прохождения более половины кадра. Преимуществом этого способа заключается в простоте конструкции, быстрота процесса и простоте использования. Модели и фон не фиксированы, они меняются со временем. Недостатками этих подходов является сбой при отслеживании целей на анимированных фонах и зависимость точности от скорости движения цели.

2) Разность кадров: Один из самых простых методов вычитания фона – это разность кадров. В этом методе предыдущий кадр считается фоном. Таким образом, цель определяется путем вычитания теку-

щих рамок фоновой модели [6-8]. Вычитывается разность кадров в момент времени $t + 1$, а фоном считается кадр в момент времени t . Это разностное изображение будет показывать только некоторую интенсивность для местоположений пикселей, которые изменились в двух кадрах. Этот подход будет работать только в случаях, когда все пиксели переднего плана движутся, а все пиксели фона статичны [3, 9].

3) Фильтрация MIN-MAX: в этом алгоритме используются три разных значения, чтобы определить какой пиксель является частью фоновой модели. Мишень показывает большую интенсивность точек на фоне и менее интенсивность на фоне в течение определенного промежутка времени [10]. Харитаоглу и др. [11] предложили другой метод с целью локальной адаптации к шуму. Здесь каждый фоновый пиксель имеет максимальную M_s , минимальную m_s и максимальную разность последовательных кадров D_s , наблюдаемую в обучающей последовательности. Большинство схем используют факторы забывания или экспоненциальное взвешивание для определения отношения вклада прошлых наблюдений, которое следует ниже. Их можно использовать для вычитания и оценки фона [10].

4) Приближенная медианная фильтрация: Макфарлейн и Шофилд [4] представили простой рекурсивный фильтр для оценки медианы каждого пикселя во времени. Этот метод был принят некоторыми подходами для вычитания фона при мониторинге городского движения из-за его значительной скорости.

5) Одиночная гауссовская фильтрация: вычисление среднего изображения последовательности кадров, а затем вычитание каждого нового входного кадра и проверка значений разности по заранее определенному порогу – один из самых простых методов вычитания фона [12]. Рен [10] представил алгоритм для распределения нормального распределения со стандартным отклонением и определенным средним значением для каждого оцениваемого фонового пикселя с использованием цветово-

го пространства, называемого цветовым пространством YUV.

6) Фильтрация Калмана. Этот метод является одним из самых известных рекурсивных методов. Если мы предположим, что значения интенсивности пикселей в изображении подчиняются нормальному распределению, например, когда простые адаптивные фильтры отвечают за обновление среднего значения и дисперсии модели фона, чтобы компенсировать изменения освещения и включить объекты с длительными остановками в модели фона. Оценка фона с использованием фильтрации Калмана была объяснена в [8]. Основное различие между ними заключается в используемом пространстве состояний для отслеживания процесса. Самыми простыми являются те, которые основаны только на яркости [3, 8, 9].

7) Скрытые марковские модели. Все упомянутые модели способны приспосабливаться к постепенным изменениям освещения. Однако, если происходят значительные изменения интенсивности, все они сталкиваются с серьезными проблемами. Метод, который может моделировать изменения интенсивности пикселей, известен как модель Маркова. Он пытается смоделировать эти вариации как дискретные состояния, основанные на режимах окружающей среды, например, облачное/солнечное небо или включение/выключение света. НММ с тремя состояниями был показан для моделирования интенсивности пикселя в приложениях мониторинга трафика [13].

Статистические методы.

Моделирование фона одним изображением, как и в обычных методах, требует строго фиксированного фона, свободного от шумов. Поскольку это требование не может быть выполнено в постоянном режиме в реальной жизни, многие модели с каждым фоновым пикселем в функции плотности вероятности (PDF – probability density function) обучались на серии обучающих кадров. Статистические методы, использующие одну гауссиану, имеют две подпоследовательности: гауссово среднее было предложено Реном [10]. Он не справляется с мультимодальными фонами.

Многие исследователи работали над статистическими методами с использованием нескольких гауссианов, которые называются гауссовской смешанной моделью (GMM – Gaussian Mixture Model). Некоторые из этих исследований были проведены Штауффером и Гримсоном [4], Живковичем [14]. Для учета фонов из анимированных текстур (таких как деревья, раскачиваемые ветром) некоторые авторы предложили использовать мультимодальные функции плотности вероятности, такие как метод Штауффера и Гримсона [15].

Непараметрические методы

Эльгаммаль и его коллеги предложили алгоритм оценки плотности ядра (KDE – Kernel Density Estimation). Неструктурированный подход также можно использовать для моделирования мультимодальной PDF. С этой точки зрения Эльгаммаль предложил оценку окна Парцена для каждого фонового пикселя. Проблема этого метода заключается в размере требования к памяти, времени для вычисления значений ядра. Также могут быть предусмотрены более сложные методы, такие как Mittal и Paragios [16], которые основаны на «ядрах с переменной пропускной способностью». Хофманн предложил адаптивный сегментатор на основе пикселей (PBAS – Pixel-Based Adaptive Segmenter).

Сравнение методов

Недавно Тиан и др. предложили метод выборочного моделирования фона и вычитания Eigen, который может сохранять надежность в многолюдных сценах. В их методы интегрированы три механизма «избирательности», включая выборочное обучение, выборочную инициализацию модели и выборочную реконструкцию на уровне пикселей. Они использовали три фоновых алгоритма Eigen: C-EigenBg, BS-EigenBg, PS-EigenBgNVF и сравнивали результаты с другими фоновыми алгоритмами, отличными от Eigen, такими как GMM, Bayes, Codebook, PBAS и Vibe.

Рай и его коллеги также представили метод сегментации на основе нейронной сети, который может извлекать движущиеся объекты из видео. Нейронная сеть объединяет входные данные, которые используют пространственно-временную корреляцию между пикселями. Показаны результаты, демонстрирующие эффективность метода по сравнению с методом MoG с более интенсивным использованием памяти, как показано на следующем рисунке. Как видно из видео, метод Luque не может эффективно сегментировать объекты переднего плана. MoG дает лучшие результаты, чем метод Luque, но метод Joudaki дает наилучшие общие результаты, как показано на рисунке 1.

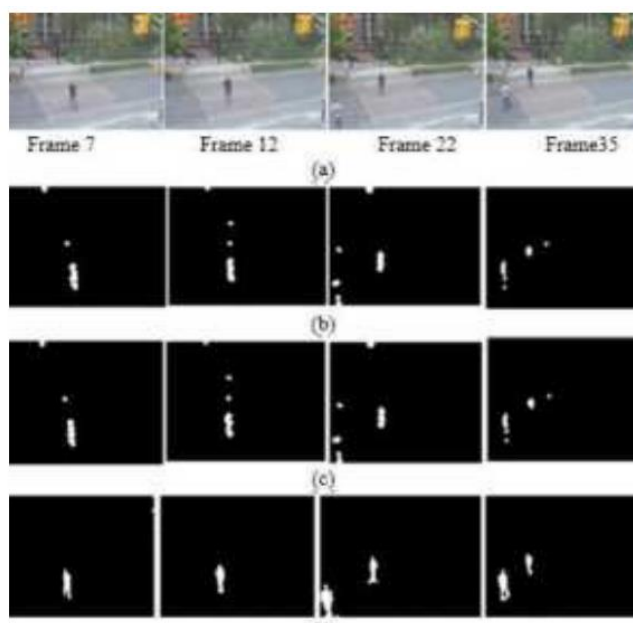


Рис. 1. а) Оригинал изображения; б) Результат метода R.M. Luque, в) Результат метода MoG, г) Результат метода Joudaki

Метод MinMax не подходит для шумных видео. Это объясняется тем, что порог MinMax (глобальный) зависит от максимальной межкадровой разницы (которая велика для видео с шумом) и, таким образом, подвержен ложным срабатываниям. Что касается основных методов, их фиксированный глобальный порог значительно снижает производительность.

Заключение. В этой статье описаны плюсы и минусы популярных методов вычитания фона. Эти методы имеют разную

структуру: от самого обычного, например Basic, до значительно сложного, например Eigen. Была исследована способность методов правильно обнаруживать объект на видео. Методы обработки видео в градациях серого, например MinMax, дают менее точный результат, чем на цветном видео. Методы, такие как KDE и GMM, дали лучший результат при большом уровне шума. Eigen потребляет огромное количество памяти, поэтому данный метод не подходит для использования в реальности.

Библиографический список

1. Патрик С., Вун Ю.В. и Комли Р. Показатели оценки производительности для отслеживания видео // Технический обзор IETE. 2011. – 493 с.
2. Тамерсой Б. Вычитание фона // Конспект лекций. 2009.
3. Куккиара Р., Грана К., Прати А., Веццани Р. Вероятностная классификация поз для анализа человеческого поведения, транзакций в системах // Человек и кибернетика. 2005. С. 42-54.
4. Макфарлейн Н.Дж.Б., Шофилд К.П. Сегментация и отслеживание порослят на изображениях // Машинное зрение и приложения. 1995. С. 187-193.
5. Кальдерара С., Прати А., Куккиара Р. Надежный фон // Доп. для сложных сцен. 2006.
6. Приложения, Третий IEEE Int. Практикум по визуальному наблюдению. 2000. С. 37-45.
7. Нараги С. Сравнительное исследование алгоритмов фоновой оценки, представленное в Институт последипломного образования и исследований в рамках частичного выполнения требований для получения степени магистра. Диплом, 2009.
8. Халеви Г., Вайншолл Д. Движение возмущений: обнаружение и отслеживание нежесткого движения нескольких тел // Maching Vision and Applications. 1999. С. 122-137.
9. Пиккарди М. Методы фонового вычитания: обзор, Международная конференция IEEE по системам // Человек и кибернетика. – 2004. – №4. – С. 3099-3104.
10. Понг П.К.Т. и Боуден Р. Улучшенная адаптивная модель смешения фона для отслеживания в реальном времени с обнаружением теней // Конференция «Семинар по передовым системам видеонаблюдения». – 2001.
11. Харитаоглу И., Харвуд Д., Дэвис Л.С. W4: Наблюдение за людьми и их действиями в режиме реального времени, анализ закономерностей и машинный интеллект. – 2000. – С. 809-830.
12. Мохаммад А.А.Х., Осман М. Техника вычитания фона адаптивного медианного фильтра с использованием нечеткой логики // Международная конференция по вычислительной технике, электротехнике и электронной технике (ICSEEE). – 2013. – С. 115-120.
13. Чинг С., Ченг С. и Камат К. Надежные методы вычитания фона в видеозаписи городского трафика // Визуальные коммуникации и обработка изображений. – 2004.
14. Живкович З. Улучшенная адаптивная гауссовская смешанная модель для вычитания фона // Международная конференция по распознаванию образов. – 2004.
15. Штауффер К., Гримсон В.Э.Л. Адаптивные смешанные модели фона для отслеживания в реальном времени // Международная конференция по компьютерному зрению и распознаванию образов. – 1999.
16. Миттал А., Парагиос Н. Вычитание фона на основе движения с использованием адаптивной оценки плотности ядра // Материалы международной конференции по компьютерному зрению и распознаванию образов. – 2004.

BACKGROUND SUBTRACTION METHODS FOR OBJECT DETECTION IN A VIDEO

A.I. Baimukhametova, *Graduate Student*

**Kazan National Research Technical University named after A.N. Tupolev – KAI
(Russia, Kazan)**

***Abstract.** The article discusses popular background subtraction methods for detecting an object in a video. Background subtraction is one of the most important parts in video processing. When processing video, it is necessary to select only those objects that have a certain value and remove unnecessary objects. The deletion is because processing video with extra items leads to an increase in processing time and required memory. This article presents a comparative study of several existing background subtraction methods, from simple background subtraction to more sophisticated statistical methods. The purpose of this study is to find the advantages and disadvantages of the methods that are widely used in the modern world.*

***Keywords:** background subtraction, video processing, computer vision, object detection.*