

**Система мультикамерного отслеживания объекта**

© А.В. Бобков, Г.В. Тедеев

МГТУ им. Н.Э. Баумана, Москва, 105005, Россия

*Предложена система мультикамерного отслеживания объекта, реализованная с помощью технологий компьютерного зрения и позволяющая оператору видеонаблюдения в режиме реального времени выбрать объект, за которым в дальнейшем будет следить система — она будет готова выдать местоположение объекта в любой момент времени. Решение данной задачи разбивается на три основных этапа: этап обнаружения, этап сопровождения и этап взаимодействия нескольких камер видеонаблюдения. Исследованы методы обнаружения, сопровождения объектов и взаимодействия нескольких камер. Для решения задачи обнаружения исследованы метод оптического потока и метод удаления заднего фона, для решения задачи сопровождения — метод сопоставления ключевых точек и корреляционный метод, для решения задачи взаимодействия нескольких камер видеонаблюдения — метод топологического графа сети камер. Предложен подход построения системы, использующий комбинацию метода удаления заднего фона, корреляционного метода и метода топологического графа сети камер. Экспериментально реализованы этапы обнаружения и сопровождения, т. е. решена задача отслеживания объекта в пределах зоны действия одной видеокamеры. Реализованная система показала хорошие результаты: достаточно высокая скорость и точность с редкими потерями сопровождаемого объекта и с небольшим уменьшением частоты кадров.*

**Ключевые слова:** обнаружение, сопровождение, метод вычитания заднего фона, корреляционный метод, мультикамерное отслеживание

**Введение.** Развитие современных технологий затрагивает многие аспекты жизни человека, в том числе его безопасность. Во все времена обеспечение безопасности было одной из основных задач, решаемых человечеством. Однако в настоящее время в связи с усложнением военно-политической, криминальной и террористической обстановки в мире для решения этой проблемы требуется более инновационный подход. Одна из главных задач гарантии безопасности — защита предприятий, зданий, помещений, а также людей, находящихся в них, от нападений, грабежа, террористических актов и т. д.

Рассмотренный в статье комплексный подход отслеживания объекта реализован с помощью технологий компьютерного зрения. Оператор видеонаблюдения в режиме реального времени может выбрать конкретный «подозрительный» объект, за которым система в дальнейшем будет следить и выдавать местоположение объекта в любой момент времени.

Цель предложенной системы — облегчить работу оператора видеонаблюдения в сфере безопасности и, следовательно, повысить степень защиты предприятий, зданий и помещений.

Решение проблемы разбивается на три главные подзадачи: 1) обнаружение объекта; 2) его сопровождение; 3) взаимодействие нескольких камер видеонаблюдения. В данной статье рассмотрены различные методы их решения и предложены оптимальные варианты реализации системы.

**Предлагаемые подходы к решению основных подзадач.** Первая задача комплексного решения — обнаружение движущихся объектов, заключающееся в выделении в видеопотоке движущихся объектов, представляющих интерес.

Существует множество методов обнаружения, из которых будут рассмотрены два основных:

- метод оптического потока [1];
- метод вычитания фона [2].

Метод оптического потока основан на понятии оптического потока, которое представляет карту сдвигов каждой точки между двумя соседними изображениями в видеопотоке. Поскольку сдвиг с точностью до масштаба эквивалентен мгновенной скорости, можно считать оптическим потоком изображение поля скоростей точек. Преимущество данного подхода — высокая точность, недостаток — высокая вычислительная требовательность, из-за которой его нецелесообразно применять в системах реального времени.

Метод вычитания фона подразделяется на метод разности кадров и метод гауссовой смеси [2]. Метод разности кадров основан на вычитании текущего кадра из фонового. Если разность в значениях каждого пиксела меньше некоторого порогового значения, то он определяется как задний фон, иначе — как передний план. Данный метод обладает высокой точностью и прост в реализации, но имеет следующие недостатки: высокую чувствительность к теням и освещению, низкую точность, а также для него требуется постоянный фон.

Метод гауссовой смеси основан на адаптивном статическом моделировании интенсивности пикселов. Выбирается смесь из нескольких гауссиан, которая используется для построения модели заднего фона. Этот метод отличается высокой точностью, достаточно большой скоростью, но чувствителен к шумам в видеоряде.

На этапе обнаружения выявляется бинарный видеопоток, к которому также применяются различные морфологические операции для снижения шумов [3].

**Сопровождение объекта.** Вторая задача отслеживания объекта — обеспечение сопровождения. Она заключается в построении траектории движения выбранных объектов на последовательности кадров видеопотока. Рассмотрим два основных метода ее решения:

- метод сопоставления ключевых точек [4];
- корреляционный метод [5].

Метод сопоставления ключевых точек основан на выделении в изображении некоторых локальных особенностей — особых точек, окрестность которых можно отличить от окрестности любой другой точки изображения [6]. В качестве окрестности обычно выбирается небольшая область пикселей вокруг точки размером примерно  $5 \times 5$ . Метод состоит из двух процессов: детектирования и дескрипирования. В процессе детектирования обнаруживаются особые точки, а в процессе дескрипирования — выделяются дескрипторы этих точек на последовательности кадров в видеопотоке. Оба процесса являются инвариантными. Данный метод обладает высокой точностью и нечувствительностью к геометрическим искажениям, но имеет высокую чувствительность к шумам, его применение должно отвечать высоким вычислительным требованиям, а также с помощью этого метода сложно определить нужное количество особых точек.

Корреляционный метод основан на корреляционном анализе изображений, что дает возможность установить соответствие между двумя изображениями и определить степень их сходства путем вычисления корреляционной функции

$$R(x, y) = \sum_{y'=0}^{h-1} \sum_{x'=0}^{w-1} T(x', y') I(x+x', y+y'), \quad (1)$$

где  $T(x', y')$  — яркость пиксела  $(x', y')$  в изображении шаблона (объекта)  $T$ ;  $I(x+x', y+y')$  — яркость пиксела  $(x+x', y+y')$  в основном изображении.

Координаты объекта на кадре видеопотока соответствуют максимуму двумерной взаимно корреляционной функции изображения объекта и всего кадра. Однако для расчета по формуле (1) требуется огромное количество вычислений. Для того чтобы разрешить данную проблему, можно использовать теорему о спектре свертки, которая заключается в том, что вместо длинного расчета кольцевой свертки в спектральной области можно применить более короткую и быструю операцию — почленное перемножение элементов спектра:

$$F(I \otimes T) = F(I) \times F^*(T), \quad (2)$$

где  $F(\ )$  — дискретное преобразование Фурье ( $\otimes$  — операция кольцевой свертки);  $*$  — операция комплексного сопряжения.

Корреляционную функцию  $R(I, T)$  можно представить в виде кольцевой свертки, дополнив нулями меньшее изображение (кадр) до большего. Очевидно, что значение корреляционной функции от этого не изменится, и тогда расчетная формула примет окончательный вид

$$R(I, T) = F^{-1}[F(I) \times F^*(T)], \quad (3)$$

где  $F^{-1}[\ ]$  — обратное преобразование Фурье.

Использование формулы (3) позволяет уменьшить объем вычислений на четыре-пять порядков. Для снижения чувствительности к условиям освещенности применяют нормирование корреляционной функции.

Данный метод обладает высокими скоростью и нечувствительностью к шуму и изменению яркости, но чувствителен к геометрическим искажениям объекта.

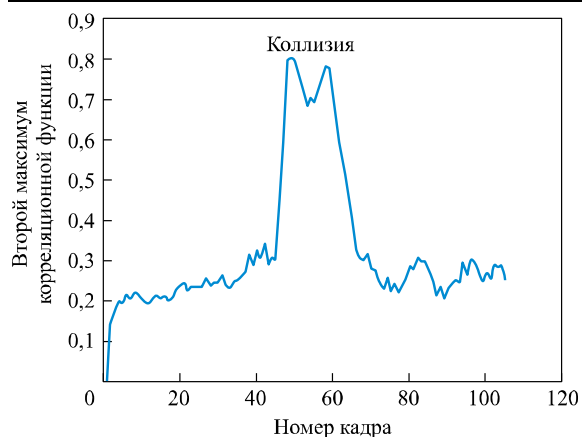
**Детектирование коллизий.** Практически в любой системе отслеживания объектов одним из критических моментов является детектирование коллизий — перекрытий объекта интереса другим подвижным либо неподвижным объектом сцены [7, 8]. В этом случае объект интереса на время исчезает из поля зрения, и тогда алгоритм отслеживания может ошибочно принять за исчезнувший объект какой-то другой объект в поле зрения, например сам объект перекрытия. Для обработки коллизии необходимо решить три задачи:

- обнаружение факта коллизии;
- замораживание образца объекта интереса;
- прогноз времени и места повторного появления объекта на сцене.

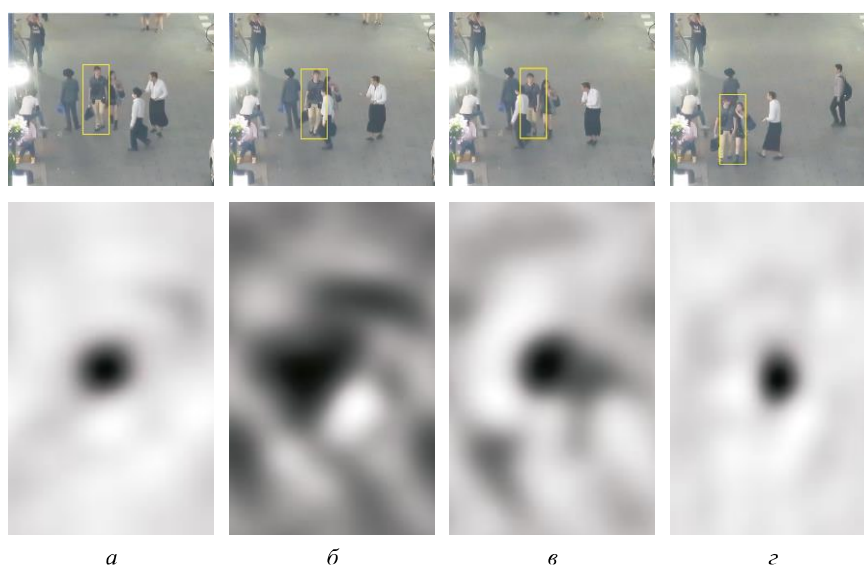
Обнаружение факта коллизии может быть основано как на анализе поведения соседних объектов сцены, так и на анализе параметров самого объекта, например формы корреляционной функции. В первом случае в окрестности объекта интереса обнаруживаются области движения, определяется их скорость и оценивается вероятность перекрытия с ними. Недостаток такого подхода заключается в том, что он работает только для изолированных областей движения. Для нескольких близкорасположенных объектов, например людей в толпе, область движения будет общей, и в этом случае выделить изолированный объект коллизии крайне проблематично.

В представленной работе был использован альтернативный подход — отслеживание формы корреляционной функции объекта интереса и его образца. При отсутствии коллизий корреляционная функция представляет собой гауссоиду с одним ярко выраженным максимумом. При возникновении коллизии форма корреляционной функции резко меняется — возникают дополнительные максимумы, отражающие наличие объекта коллизии. Соотношение величин главного и второстепенного максимума можно рассматривать как критерий коллизии: если он выше порога, то регистрируется коллизия, обновление образца прекращается, а прогноз положения строится по ранее измеренной скорости (рис. 1). Возврат отношения максимумов к допороговому значению свидетельствует об окончании коллизии (рис. 2).

Для стационарных объектов коллизии дополнительно могут быть размечены области потери объектов: при попадании в данную область обновление образца отключается, отслеживание выполняется по прогнозу скорости и при повторном обнаружении объекта в области прогноза его отслеживание возобновляется.



**Рис. 1.** Детектирование коллизий с использованием величины второго максимума корреляционной функции



**Рис. 2.** Форма корреляционной функции:

*а* — до коллизии; *б* — момент начала коллизии;  
*в* — момент окончания коллизии; *г* — после коллизии

При реализации межкамерного взаимодействия аналогичным образом фиксируется изменение формы корреляционной функции, а также проверяется нахождение объекта в зоне выхода. При этом фиксируется выход объекта из поля зрения камеры и инициируется его поиск на связанной камере в зоне входа. Ожидаемое время появления вычисляется исходя из скорости объекта и среднего времени перемещения объектов между данной парой входов и выходов.

Если невозможно повторно обнаружить объект в течение выделенного момента времени, то объект считается потерянным. В этом случае необходимо известить оператора, для того чтобы он выполнил

дальнейшие действия. При исчезновении объекта вблизи границ кадра, не связанных ни с одной из зон выхода, объект считается покинувшим зону наблюдения.

**Взаимодействие нескольких видеокамер.** Задача взаимодействия нескольких камер видеонаблюдения сводится к построению топологического графа сети камер. Узлами данного графа являются некоторые области видимости камер, где объект может либо появиться, либо пропасть. Ребра графа определяют реальные связи областей входа-выхода между камерами видеонаблюдения. Веса ребер указывают на примерное время перехода объекта из одного узла в другой, что повышает точность отслеживания [9].

Для построения данного топологического графа можно применить два метода: вероятностный [7] и корреляционный [10].

Вероятностный метод использует окна Парзена (также известные как оценки плотности ядра) для оценки вероятностей межкамерного пространства-времени на основе обучающих данных, т. е. вероятности попадания объекта в определенную камеру в определенное время с учетом местоположения, времени и скорости его выхода с других камер. Вероятность соответствия, т. е. вероятность того, что два наблюдения относятся к одному и тому же объекту, зависит как от пространственно-временной информации, так и от внешнего вида. Преимущество данного метода — высокая точность, но для него требуются начальные обучающие данные.

Корреляционный метод за событиями входа-выхода объектов наблюдения измеряет корреляции между событиями, для того чтобы установить топологию сети камер. Этот метод целесообразно применять на практике, поскольку не требуется выполнение таких шагов по согласованию внешнего вида, как повторная идентификация и отслеживание между камерами для вывода топологии. Однако топология, предполагаемая основанная на событиях подходом, может оказаться неточной, так как может быть выведена из ложных событий — появления или пропажи объекта из зоны наблюдения. Кроме того, в случае применения этого подхода требуется выполнить значительный объем вычислений при большом количестве камер в сети видеонаблюдения.

Недостаток статистических подходов заключается в том, что возможно существование редко используемых путей, для которых не удастся набрать достаточного количества статистической информации. В этом случае система просто не сумеет найти такую пару входа-выхода, попав в которую объект интереса будет бесследно пропадать и его будет невозможно повторно обнаружить. Следует отметить, что особенно большую проблему это будет представлять в закрытых помещениях с ограниченным числом проходов между зонами наблюдения. К тому же ручная разметка в данном случае

затруднена из-за того, что необходимо задавать не только пространственные связи между областями, но и временные. Поэтому в рассматриваемых случаях был дополнительно введен режим настройки межкамерного взаимодействия, который работает в совершенно пустом помещении, например, непосредственно после установки и подключения камер.

В помещении перемещается оператор-настройщик, последовательно проходя все известные ему проходы между помещениями. Он одет в контрастную одежду (оранжевую куртку), которая служит маркером для системы межкамерного взаимодействия.

Для обнаружения маркера система находит области движения, фильтрует и размечает их, а затем выделяет среди них наиболее крупную область, имеющую заданные тон и яркость. Фильтрация по движению позволяет отделить маркер от стационарных объектов того же цвета, а фильтрация по тону — убрать ложные области движения, возникающие, например, из-за теней и бликов.

Таким образом, система работает с единственным объектом интереса, который однозначно идентифицируется как настройщик на всех кадрах всех камер. Фиксируя моменты исчезновения и появления оператора между кадрами, система строит пары зон исчезновения и появления объекта, измеряет скорость движения, время между исчезновением и появлением, а также оценивает виртуальное «расстояние» в пикселях между парой зон входа-выхода (рис. 3).



**Рис. 3.** Пример топологического графа межкамерного взаимодействия:  
1 — зона выхода из поля наблюдения камеры I; 2 — зона выхода из поля камеры I в камеру II; 3 — зона входа из поля камеры I в камеру II; 4 — зона выхода из поля наблюдения камеры II

Затраты на реализацию режима настройки при относительно небольшом количестве камер оказываются несущественными, а точность построения топологии областей видеонаблюдения значительно возрастает, поэтому данный режим выбран в качестве основного.

**Система мультикамерного отслеживания.** На первом этапе этой системы выделяются области движения с использованием опорного кадра и метода гауссовой смеси. Найденные точки движения с помощью морфологических методов объединяются в более крупные области, а изолированные точки считаются шумовыми и удаляются. Также удаляются области с низким градиентом яркости, поскольку они не содержат объекты и являются артефактами. Затем происходит разметка найденных движущихся областей.

Далее оператор может выбрать любую движущуюся область и пометить ее как подлежащую отслеживанию. Оператор может также задать область отслеживания вручную, если она была почему-либо пропущена детектором движения. При выборе области система переходит в режим отслеживания объекта интереса. Для отслеживания объекта был использован модифицированный трекер МОССЕ. На каждом очередном шаге трекер обновляет образец объекта интереса в спектральной области, обеспечивая устойчивость к изменению формы объекта. Это необходимо, потому что фигура человека сильно меняется в процессе движения, а значит, одного первоначального образца для отслеживания здесь недостаточно.

В процессе отслеживания также контролируется форма корреляционной функции. При появлении дополнительных существенных максимумов детектируется факт коллизии, в связи с чем отслеживание переходит в режим прогноза движения до возвращения корреляционной функции к исходной нормальной форме.

Если в момент коллизии объект интереса попал в зону выхода, фиксируется межкамерная коллизия, определяются связанная камера и область входа, рассчитывается примерное время появления в ней объекта. Появление объекта в области выхода фиксируется с пониженным порогом, поскольку ракурс съемки объекта на входе и на выходе может не совпадать.

Система была реализована на языке Python с использованием открытой библиотеки компьютерного зрения OpenCV. Общий алгоритм представлен на рис. 4. На рис. 5–8 приведены основные этапы работы системы. Реализованная система показала достаточно высокие скорость и точность, наблюдаются лишь редкие потери сопровождаемого объекта и небольшая задержка в частоте кадров. Для проверки работоспособности системы использовались как известные размеченные видеонаборы (Multiple Object Tracking Benchmark-16), так и собственное видео. Эксперименты продемонстрировали, что система способна работать в режиме реального времени, отслеживая объекты на изображении размером 1920×1080 пикселей со скоростью не менее 24 кадр/с. Этого достаточно для работы системы в режиме реального времени, а также, потенциально, — для отслеживания сразу нескольких объектов.



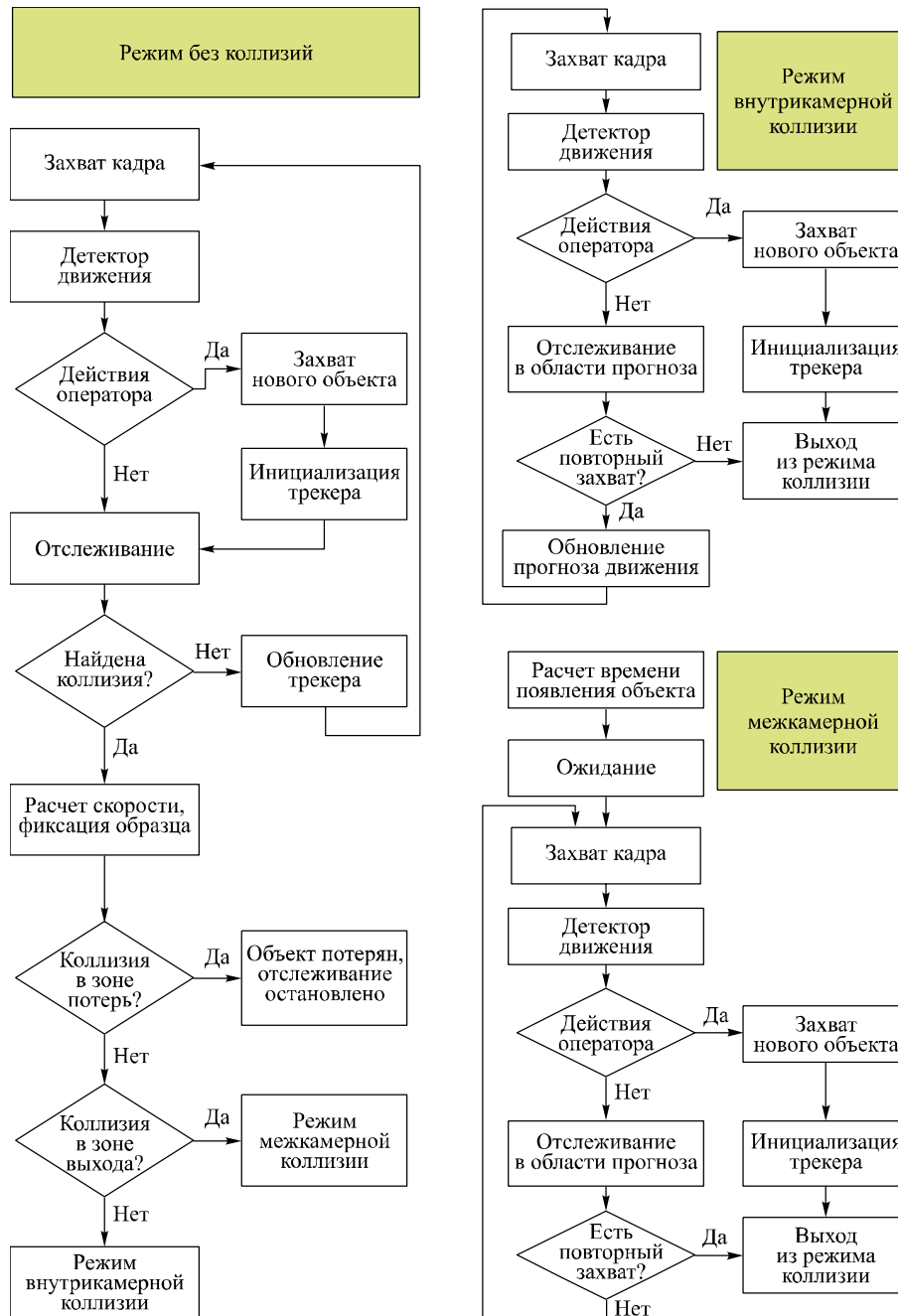


Рис. 4. Алгоритм работы системы обнаружения и сопровождения

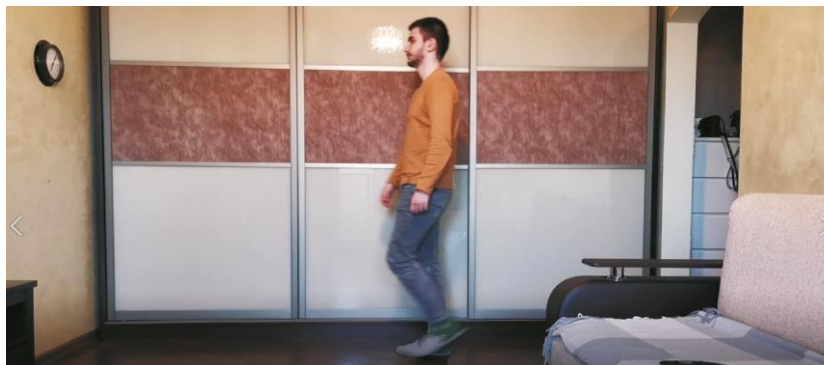


Рис. 5. Входной кадр видеопотока



Рис. 6. Бинарная маска кадра

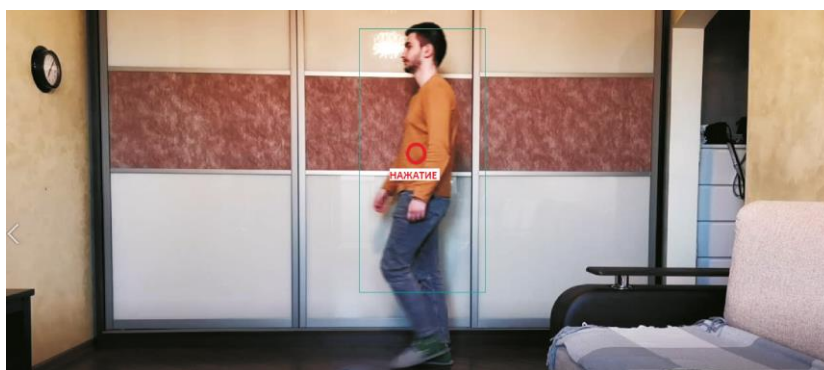


Рис. 7. Детектирование нажатия оператора и инициализация трекера

Система также показала возможность отслеживать объекты при наличии коллизий. Вероятность потери объекта при коллизии как внутри кадра, так и при межкамерном взаимодействии составила менее 5 %, вероятность потери объекта вне коллизии оказалась менее 1 %. В основном потери происходили при движении в толпе, когда выделить область движения отдельного человека не представлялось возможным.



Рис. 8. Сопровождение выбранного объекта

**Заключение.** В данной статье рассмотрен подход к проектированию системы мультикамерного отслеживания. Задача была разбита на три основные подзадачи: отслеживание, сопровождение и взаимодействие нескольких видеокамер. Для решения задачи обнаружения исследованы метод оптического потока и метод вычитания фона (метод гауссовой смеси). Для решения метода сопровождения рассмотрены метод сопоставления ключевых точек и корреляционный метод. Для решения взаимодействия камер видеонаблюдения использовался метод построения топологического графа сети камер, а для решения задачи его построения представлены вероятностный метод и корреляционный метод. В предложенном подходе проектирования системы отслеживания были выбраны метод вычитания фона (метод гауссовой смеси), корреляционный метод и метод построения топологического графа сети камер с помощью кросс-корреляции. Реализованная система показала хорошие результаты: достаточно высокую скорость и высокую точность с редкими потерями сопровождаемого объекта и с небольшим уменьшением частоты кадра.

## ЛИТЕРАТУРА

- [1] Agarwal A.A., Gupta S., Singh D.K. *Review of optical flow technique for moving object detection*. IEEE, 2016, pp. 409–413.
- [2] Заливин А.Н., Балабанова Н.С. Обнаружение движущихся объектов методом вычитания фона с использованием смеси Гауссовых распределений. *Автоматизированные технологии и производства*, 2016, № 3, с. 29–32.
- [3] Берников В.В., Преображенский А.П., Чопоров О.Н. Анализ алгоритмов обнаружения движущихся объектов на видеоизображении. *Моделирование, оптимизация и информационные технологии*, 2018, № 6 (3), с. 223–233.
- [4] Kloihofe W., Kampel M. Interest point based tracking. *IEEE*, 2010, pp. 3549–3552.
- [5] Bolme D.S., Beveridge J.R., Draper B.A., Lui Y.M. Visual object tracking using adaptive correlation filters. *IEEE*, 2010, pp. 2544–2550.

- [6] Kustikova V.D., Gergely V.P. Vehicle video detection and tracking quality analysis. *Pattern Recognition and Image Analysis. Advances in Mathematical Theory and Applications*, 2016, vol. 26, no. 1, pp. 108–116.
- [7] Javed O., Rasheed Z., Shafique K., Shah M. Tracking across multiple cameras with disjoint views. *IEEE*, 2003, 952 p.
- [8] Wang Y., Lu K., Zhai R. Challenge of multi-camera tracking. In: *2014 7th International Congress on Image and Signal Processing*. IEEE, 2014, October, pp. 32–37.
- [9] Hatwar R.B., Kamble S.D., Thakur N.V., Kakde S. A review on moving object detection and tracking methods in video. *International Journal of Pure and Applied Mathematics*, 2018, no. 118 (16), pp. 511–526.
- [10] Makris D., Ellis T., Black J. Bridging the gaps between cameras. *IEEE*, 2004, vol. 2, pp. II-205–II-210.

Статья поступила в редакцию 19.05.2021

Ссылку на эту статью просим оформлять следующим образом:

Бобков А.В., Тедеев Г.В. Система мультикамерного отслеживания объекта.

*Инженерный журнал: наука и инновации*, 2021, вып. 10.

<http://dx.doi.org/10.18698/2308-6033-2021-10-2123>

**Бобков Александр Валентинович** — канд. техн. наук, доцент кафедры «Системы автоматического управления» МГТУ им. Н.Э. Баумана. e-mail: [skliz@mail.ru](mailto:skliz@mail.ru)

**Тедеев Георгий Вадимович** — бакалавр кафедры «Системы автоматического управления» МГТУ им. Н.Э. Баумана. e-mail: [georgetd@mail.ru](mailto:georgetd@mail.ru)

## Multi-camera object tracking system

© A.V. Bobkov, G.V. Tedeev

Bauman Moscow State Technical University, Moscow, 105005, Russia

*The article proposes a multi-camera tracking system for an object, implemented using computer vision technologies and allowing the video surveillance operator in real time to select an object that will be monitored by the system in future. It will be ready to give out the location of the object at any time. The solution to this problem is divided into three main stages: the detection stage, the tracking stage and the stage of interaction of several cameras. Methods of detection, tracking of objects and the interaction of several cameras have been investigated. To solve the problem of detection, the method of optical flow and the method of removing the background were investigated, to solve the problem of tracking — the method of matching key points and the correlation method, to solve the problem of interaction between several surveillance cameras — the method of the topological graph of a network of cameras. An approach is proposed for constructing a system that uses a combination of the background removal method, the correlation method and the method of the topological graph of a network of cameras. The stages of detection and tracking have been experimentally implemented, that is, the task of tracking an object within the coverage area of one video camera has been solved. The implemented system showed good results: a sufficiently high speed and accuracy with rare losses of the tracked object and with a slight decrease in the frame rate.*

**Keywords:** detection, tracking, background subtraction, correlation method, multi-camera tracking

### REFERENCES

- [1] Agarwal A.A., Gupta S., Singh D.K. Review of optical flow technique for moving object detection. *IEEE*, 2016, pp. 409–413.
- [2] Zalivin A.N., Balabanova N.S. *Avtomatizirovannye tekhnologii i proizvodstva — Automation of technologies and production*, 2016, no. 3, pp. 29–32.
- [3] Bernikov V.V., Preobrazhenskij A.P., Choporov O.N. *Modelirovanie, optimizaciya i informacionnye tekhnologii — Modeling, Optimization and Information Technology*, 2018, vol. 6 (3), pp. 223–233.
- [4] Klohofer W., Koppel M. Interest point based tracking. *IEEE*, 2010, pp. 3549–3552.
- [5] Bolme D.S., Beveridge J.R., Draper B.A., Lui Y.M. Visual object tracking using adaptive correlation filters. *IEEE*, 2010, pp. 2544–2550.
- [6] Kustikova V.D., Gergely V.P. Vehicle video detection and tracking quality analysis. *Pattern Recognition and Image Analysis. Advances in Mathematical Theory and Applications*, 2016, vol. 26, no. 1, pp. 108–116.
- [7] Javed O., Rasheed Z., Shafique K., Shah M. Tracking across multiple cameras with disjoint views. *IEEE*, 2003, 952 p.
- [8] Wang Y., Lu K., Zhai R. Challenge of multi-camera tracking. In: *2014 7th International Congress on Image and Signal Processing*. IEEE, 2014, pp. 32–37.
- [9] Hatwar R.B., Kamble S.D., Thakur N.V., Kakde S. A review on moving object detection and tracking methods in video. *International Journal of Pure and Applied Mathematics*, 2018, vol. 118 (16), pp. 511–526.

- [10] Makris D., Ellis T., Black J. Bridging the gaps between cameras. *IEEE*, 2004, vol. 2, pp. II-205–II-210.

**Bobkov A.V.**, Cand. Sc. (Eng.), Assoc. Professor, Department of Automatic Control Systems, Bauman Moscow State Technical University. e-mail: skliz@mail.ru

**Tedeev G.V.**, bachelor, Department of Automatic Control Systems, Bauman Moscow State Technical University. e-mail: georgetd@mail.ru