

АЛГОРИТМ ОТСЛЕЖИВАНИЯ ПЕРЕМЕЩЕНИЙ ЧЕЛОВЕКА В ВИДЕОПОТОКЕ С КАМЕР ВИДЕОНАБЛЮДЕНИЯ

© 2023 М. М. Гуменюк✉, А. В. Бровко

*Саратовский государственный технический университет им. Гагарина Ю. А.
ул. Политехническая, 77, 410054 Саратов, Российская Федерация*

Аннотация. Задача отслеживания перемещений объектов в видеопотоке (трассировка) является одной из наиболее сложных задач современных систем видеонаблюдения и видеоаналитики. Под задачей трассировки понимается задача автоматизированного распознавания и дальнейшего отслеживания объектов на протяжении последовательности кадров видеопотока. Целью данной статьи является не только краткий анализ некоторых существующих методов и алгоритмов, необходимых для визуализации движения объектов, но и разработка нового алгоритма трассировки перемещений человека в видеопотоке с возможностью отображения траекторий движения. Предложенный подход позволяет устранить некоторые недостатки существующих в настоящее время методов (наличие специального оборудования, необходимость предварительного обучения, перекрытие объекта другими объектами, выход за рамки поля зрения камеры). В работе описаны основные преимущества данного алгоритма, его важнейшие функции и возможности. Основные этапы работы алгоритма включают в себя: выбор в видеопотоке человека, осуществление процесса цифровой обработки, основой которого является разделение тела человека на составляющие части и получение цветовой гистограммы этих частей, прогнозирование локализации и распознавание исследуемого человека на всех последующих кадрах видеопотока, где осуществляется анализ цветových пикселей RGB на основе методики сопоставления полученных данных с цветовой гистограммой составляющих частей основного объекта. Выходные данные предложенного алгоритма применяются в процедуре формирования и отображения общей картины перемещения конкретного человека в рамках всего видеопотока. Данная статья может быть интересна специалистам и экспертам, применяющим в своих работах методы компьютерного зрения для автоматизированного получения необходимых данных при анализе фрагментов видео.

Ключевые слова: камеры наблюдения, распознавание образов, дельта E, трассировка, нейронная сеть YOLO, активная область, обнаружение объектов, трекинг, математическая статистика, прогнозируемая область, пиксели RGB.

ВВЕДЕНИЕ

Актуальность темы

В современном мире системы видеонаблюдения являются одной из самых быстроразвивающихся и перспективных сфер в области информационных технологий и играют существенную роль в сборе и анализе различных видов информации. Системы видеонаблюдения

в последнее время сформировались в отдельный сегмент средств обеспечения безопасности, как на частной территории, так и на общественных и производственных объектах. Ими оснащается большинство предприятий, организаций и различных объектов инфраструктуры с целью визуального контроля за сотрудниками, посетителями и служебными помещениями для своевременного реагирования на несанкционированные действия.

Одной из ключевых задач систем видеонаблюдения является трассировка перемещений различных объектов (в частности

✉ Гуменюк Михаил Михайлович
e-mail: GumenRoma@mail.ru



Контент доступен под лицензией Creative Commons Attribution 4.0 License.
The content is available under Creative Commons Attribution 4.0 License.

человека). Она является основой множества интеллектуальных систем компьютерного зрения и представляет собой систему отслеживания перемещений людей путем локализации их позиций на каждом кадре в рамках всего видеопотока. Информация, полученная в результате применения методов и алгоритмов трассировки объекта, может быть использована для анализа основных характеристик объектов (размера, скорости, направления движения, пути объекта) или для прогнозирования дальнейшего перемещения объектов. Данный процесс отслеживания перемещений объектов применяется во многих задачах видеоаналитики, среди которых наиболее актуальными на сегодняшний день являются:

- мониторинг дорожных ситуаций;
- процесс вычисления количества сотрудников на промышленных предприятиях при входе/выходе и въезде/выезде с территории с возможностью распознавания лиц и номерных знаков транспортных средств;
- мониторинг действий персонала и посетителей в торговых комплексах и других общественных местах;
- анализ работы сотрудников государственных предприятий и организаций и частных компаний.

Для более полной характеристики рассматриваемого вопроса был проведен анализ работ современных исследователей (Т. Анштедт, И. Келлер, Х. Лутц, П. Виола, Е. Ю. Минаев, В. В. Кутикова, А. А. Артемов, М. В. Кавалеров, С. Д. Шелабин, Я. Э. Живрин, А. В. Пчелкина и др.) по данной тематике, который демонстрирует тот факт, что на сегодняшний день задача трассировки объектов не является полностью решенной. Усиление внимания к проблеме сопровождения объектов в видеопотоке связано с наличием определенных сложностей, среди которых следует выделить следующие: изменения условий освещенности, перекрытие объектов другими объектами, проективные и аффинные искажения и другие [1–7].

Существующие методы и их недостатки

Теоретический анализ литературы позволяет выделить некоторые существующие современные методы визуализации движения объектов в видеопотоке. К примеру,

- Online Boosting Tracker [8] — обладает высокой точностью отслеживания. Данный метод основан на алгоритме AdaBoost с использованием каскада Хаара для поиска;
- MIL Tracker [9] — представляет собой улучшенную версию Online Boosting Tracker с применением идеи обучения с учителем;
- KCF Tracker [10] — представляет собой объединение двух вышеперечисленных алгоритмов, отличается более простой реализацией и увеличением быстродействия поиска объектов.

Необходимо отметить, что перечисленные выше методы обладают крайне важным недостатком для отслеживания перемещений объектов — отсутствием возможности повторного поиска объекта при его потере в видеокдрах (перекрытие другими объектами или выход за рамки поля зрения камер).

На сегодняшний день для устранения данной проблемы были разработаны такие методы, как Google Class, Kinect SDK, Vuforia SDK, OpenCV, но они обладают другими недостатками — необходимостью наличия специального оборудования и длительностью процесса предварительного обучения [11, 12].

Предлагаемый способ решения задач

Предлагаемый автором алгоритм базируется на основе цифровой обработки исследуемого человека. Особенностью этой стратегии является разбиение тела человека на три составляющие части, каждая из которых подвергается процессу получения их уникальной цветовой гистограммы.

Подобная процедура решения задачи отслеживания перемещений людей в видеопотоке позволяет устранять следующие проблемы:

- 1) необходимость использования специального оборудования;
- 2) необходимость предварительного обучения;

- 3) потеря исследуемого человека при перекрытии его другими объектами;
- 4) потеря человека при его выходе за пределы поля зрения камер.

1. ОПИСАНИЕ ПРЕДЛАГАЕМОГО АЛГОРИТМА

Данный алгоритм является весьма сложным и состоит из множества взаимосвязанных этапов. На рис. 1 представлена блок-схема, которая включает в себя основные этапы разработки предлагаемого алгоритма. Для тестирования предложенного алгоритма была разработана программа на языке C# в среде Visual Studio 2022.

1.1. Разделение видео на кадры

Исследуемое видео анализируется по следующим параметрам — формату видео, разрешению, длине и количеству кадров видеопотока. С целью оптимизации алгоритма был выбран способ выборочного анализа кадров — разделение видео на множество кадров с разницей в N секунд в зависимости от постановки задачи или требований пользователя. Указанный подход основывается на том, что для трассировки объектов не следует рассматривать все кадры видеопотока, а только те кадры, которые были отобраны согласно

определенным требованиям, что позволит ускорить процедуру анализа видеок кадров и уменьшить размеры хранения информации.

Для наглядности приведем пример видео, основными характеристиками которого являются — длительность 2 часа, качество 480p, вес 1 гб, формат avi, 24 кадра в секунду. В табл. 1 указано количество кадров при полном и выборочном анализе при значении $N = 1$ (один кадр в секунду). Данная таблица наглядно демонстрирует, что при выборочном анализе будут проанализированы всего 7 200 кадров вместо 172 800 кадров, что значительно ускорит работу по анализу видео.

Таким образом, пользователь получает возможность загружать интересующее его видео в алгоритм и определять, с какой периодичностью по времени следует отбирать кадры в данном видео. Алгоритм с помощью методов библиотеки OpenCV анализирует и сохраняет в специально созданную папку все кадры, которые соответствуют требованиям отбора кадров. После выполнения данной операции алгоритм приступает к следующему этапу.

1.2. Выбор одного из множества объектов

На втором этапе осуществляется выборка объекта (человека), который в дальнейшем будет подвергаться процессу цифровой обработки.

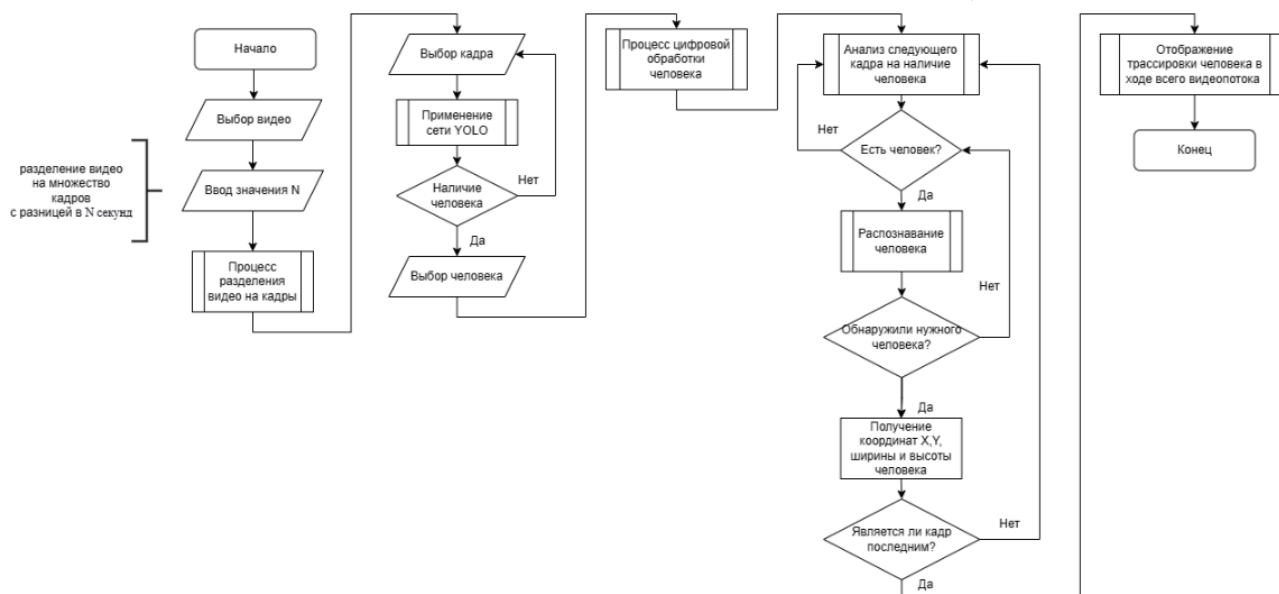


Рис. 1. Блок-схема основных этапов алгоритма
 [Fig. 1. Block diagram of the main stages of the algorithm]

Таблица 1. Сравнение количества кадров при полном и выборочном анализе
 [Table 1. Comparison of the number of frames in full and sample analysis]

	Кол-во кадров в секунду	Кол-во кадров в минуту	Кол-во кадров в час	Кол-во кадров за 2 часа (полное видео)
Полный анализ	24 кадра	1 440 кадров	86 400 кадров	172 800 кадров
Выборочный анализ	1 кадр	60 кадров	3 600 кадров	7 200 кадров

Пользователь выбирает любой интересующий его кадр из заранее созданной папки, включающей в себя все кадры, полученные во время выполнения операций первого этапа работы предложенного алгоритма. После выбора кадра, метод предоставляет пользователю два варианта выбора одного объекта из множества объектов на рассматриваемом кадре:

- 1) самостоятельное выделение объекта на кадре;
- 2) применение нейронной сети YOLO.

На сегодняшний день сверточная нейронная сеть YOLO является одной из самых часто используемых сетей для обнаружения и распознавания множества объектов на изображениях или видеокдрах. Ее отличает высокая точность получения данных об обнаруженных объектах [13–15].

Данная архитектура накладывает на входное изображение регулярную сетку $S \times S$ областей. После нахождения всех объектов на изображении сеть YOLO начинает анализировать и отбирать из них только те объекты, которые соответствуют данным класса Person сети YOLO.

Пользователю предоставляется возможность выбрать один из предложенных сетью YOLO объектов, который будет в дальнейшем применяться для прогнозирования локализации этого объекта на последующих кадрах в рамках всего видеопотока [3]. Необходимо отметить, что в качестве выходных данных вышеуказанных вариантов выступают начальные координаты, длина и ширина выделенной области, на которой зафиксирован интересующий пользователя объект (см. рис. 2). Следует обратить внимание на то, что на рис. 2 и последующих рисунках и в таблицах демонстрируются результаты работы программы, тестирующей предложенный алгоритм.

1.3. Обработка данного объекта

Третий этап является одним из ключевых этапов построения предлагаемого автором алгоритма, целью которого является получение всей цветовой информации о рассматриваемом объекте.

Первоначально возникает необходимость устранения цветowych пикселей заднего фона. Для решения этой проблемы на сегодняшний день существует интерактивный алгоритм GrabCut, способный выделять объект и удалять задний фон.

Но важнейшим недостатком этого алгоритма является вмешательство пользователя, а именно — необходимость поставить особые метки на объект (рисуют красную линию) и задний фон (отмечают синей линией), чтобы алгоритм получил возможность проанализировать, что следует вырезать, а что оставлять. Для того, чтобы алгоритм был более автоматизированным, было принято решение использовать шаблон. Шаблоном является такой кадр из видеопотока, на котором отсутствуют какие-либо объекты. Таким образом, данная стадия процесса цифровой обработки объекта состоит в сопоставлении области объекта с шаблоном, результатом которого является отделение объекта от заднего фона [16, 17].

Рассмотрим следующую стадию анализа обработки объекта. Стоит обратить внимание на то, что предлагаемый алгоритм поддерживает формат цветов — RGB. В качестве примера приведем серый цвет, который в формате RGB может содержать десятки различных значений.

К примеру, существуют два значения серого цвета (119, 136, 153) и (115, 128, 144). Для человеческого восприятия между ними прак-



Рис. 2. Выбор объекта
[Fig. 2. Object selection]

тически нет разницы, а для машинного алгоритма они представляют собой совершенно разные цвета. Для данного формата объект может содержать огромное количество различных по значению цветовых пикселей. С целью упрощения и ускорения процесса поиска объекта на последующих кадрах, был использован существующий в настоящее время алгоритм сегментации объекта — группирование объектов с близкими признаками (MeanShift). В 1975 году К. Фукунага и Л. Хостетлер впервые предложили данный алгоритм [18]. Суть этой стратегии заключается в том, что все близкие по значению и расположенные недалеко друг от друга цветовые пиксели преобразовываются в один единственный цветовой сегмент [19].

Заключительной частью обработки объекта является разделение объекта на три изображения — голову, тело и ноги. После этого для каждого из трех изображений применяется алгоритм нахождения цветового значения каждого пикселя и получения наиболее часто встречающихся цветовых пикселей.

На рис. 3 отображены все основные этапы обработки данного объекта.

1.4. Анализ каждого кадра на нахождение данного объекта

Процесс анализа каждого кадра состоит из нескольких последовательных частей:

- 1) получение прогнозируемой области изучения;
- 2) нахождение и распознавание объекта в прогнозируемой области;
- 3) получение координат объекта, запись в файл сохранения и открытие следующего кадра.

Прогнозируемая область — это некая область видеокadra, в которой предположительно может находиться нужный объект. Следует обратить внимание на то, что размер самой области зависит от расположения камер видеонаблюдения и размера объектов, а также от координат объекта с предыдущего кадра.

Данная область необходима для того, чтобы алгоритм применялся для нахождения объекта только в прогнозируемой области, а не по всему кадру, размер которого может значительно превышать размер области. Указанный процесс позволит заметно снизить ресурсы и время для обнаружения объекта (см. рис. 4).

После получения координат прогнозируемой области алгоритм на текущем кадре вырезает эту область и проводит процедуру



Рис. 3. Результат обработки данного объекта
 [Fig. 3. The result of processing this object]



Рис. 4. Получение прогнозируемой области в зависимости от размера объекта
 [Fig. 4. Getting the predicted area depending on the size of the object]

группирования объектов с близкими признаками. Вышеуказанная операция была описана в предыдущем разделе.

Следующая часть посвящена обнаружению местоположения нужного объекта с помощью цветowych пикселей в данной прогнозируемой области. Прежде чем приступать к процессу поиска соответствующих цветowych

пикселей, необходимо иметь в виду, что один и тот же объект на разных кадрах видео может отличаться из-за изменения уровня интенсивности освещения и смещения оттенков цветов. Другими словами, бессмысленно искать точно такие же значения RGB, которые были получены в процессе обработки исходного объекта.

Поэтому для решения этой проблемы было принято решение использовать формулу цветового отличия. Это математическая формула, позволяющая вычислять различие между цветовым значением исходного пикселя и цветовым значением, полученным при анализе рассматриваемого объекта. Это различие между двумя цветами называют Дельтой E (Delta E) и соответствует шкале от 0 до 100. Чем ниже показатель Дельты E, тем выше схожесть двух цветов [20].

$$\Delta E = \sqrt{(R_2 - R_1)^2 + (G_2 - G_1)^2 + (B_2 - B_1)^2}. \quad (1)$$

Представленный алгоритм работает следующим образом — если различие между цветовым значением пикселя прогнозируемой области и значением объекта соответствует $0 < \Delta E < 7$, то координаты этого пикселя записываются в специальную таблицу, которая потом будет применяться в дальнейшем. Следует обратить внимание на то, что грудь человека имеет самую большую площадь и является центром всего тела, а значит необходимо в первую очередь искать все пиксели, чьи цветовые значения согласно Дельте E соответствуют значениям, полученным при анализе груди исходного объекта.

Следующая задача — анализ всех данных из специальной таблицы. Необходимо помнить, что в прогнозируемой области могут встречаться посторонние предметы и объекты, содержащие в себе некоторые одинаковые цветовые значения. Поэтому следующим шагом является нахождение наибольшей области (области № 1), содержащей в себе ранее рассмотренные цветовые пиксели. Для этого необходимо проигнорировать все мелкие

объекты и объекты, которые располагаются на большом расстоянии друг от друга (Зависит от расположения камеры и размера прогнозируемой области).

После нахождения области № 1 необходимо вычислить минимальное и максимальное значение координат X, чтобы из прогнозируемой общей области выделить подобласть, содержащую в себе область № 1.

После нахождения области № 1, необходимо убедиться в том, что рассмотренный объект соответствует исходному объекту. Следующий шаг связан с изучением головы объекта и требует выделения прогнозируемой области над областью № 1. В этой прогнозируемой области применяется алгоритм нахождения нужных пикселей, которые состоят из соответствующих цветовых значений согласно данным из анализа головы исходного объекта, что приводит к получению области, которую назовем № 2. Аналогичным способом можно найти область № 3, содержащую в себе цветовые пиксели, соответствующие нижней части человека. Выходными данными является объединение всех трех областей и получение начальных координат, длины и ширины новой области, которые будут применяться для получения трассировки данного объекта (см. рис. 5).

Необходимо отметить, что преимуществом алгоритма является возможность сохранения всей информации об исходном объекте при анализе каждого кадра.

Если в прогнозируемой области не был обнаружен рассматриваемый объект, следовательно, данный объект вышел за пределы кадра. Но существует вероятность того, что



Рис. 5. Процесс поиска объекта в прогнозируемой области
[Fig. 5. The process of searching for an object in the predicted area]

объект может вернуться в поле зрения камеры. В этом случае на следующем кадре прогнозируемая область меняется и алгоритм выделяет все возможные области, в которых объект может оказаться при возвращении (эти области зависят от местоположения камеры и возможных вариантов входа в поле зрения камеры).

Если объект вернется в поле зрения камеры, то окажется в прогнозируемой области, где будет произведен процесс нахождения объекта. После обнаружения объекта в прогнозируемой области, прогнозируемая область на следующем кадре изменится и охватит те области, в которых объект может потенциально находиться.

После анализа всех кадров видеопотока, вся необходимая информация (координаты X и Y , ширина и высота прямоугольника) с каждого кадра записывается в заранее созданную таблицу, которая будет применяться для получения общей трассировки выбранного объекта.

2. РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

В данном разделе рассмотрим параметры, которые являются ключевыми для эффективной работы предлагаемого алгоритма. Таковыми параметрами являются:

- 1) показатель Дельты E (ΔE);
- 2) размер прогнозируемой области.

Первый этап исследования посвящен анализу результатов работы алгоритма в зависимости от различных значений Дельты E (ΔE). Полученные результаты приведены в табл. 2 и представлены в виде определенных символов. Если алгоритм правильно зафиксировал объект или его отсутствие, данные результаты считаются успешными и обозначаются символом (+). В остальных случаях отмечаются символом (-).

Данные, приведенные в табл. 2, позволяют утверждать, что при значениях $\Delta E < 7$ и $\Delta E < 12$ алгоритм демонстрирует высокую точность обнаружения объектов по сравнению с остальными значениями Дельты E .

Следующим этапом исследования является поиск наиболее оптимального значения

Дельты E с точки зрения производительности работы алгоритма. Для этого необходимо вычислить время процесса анализа каждого кадра и получить среднее значение времени в зависимости от значений Дельты E .

Согласно данным, представленным в табл. 2 и табл. 3, наиболее оптимальным значением Дельты E с точки зрения точности и производительности является $\Delta E < 7$.

Следующим этапом исследования является определение оптимального размера прогнозируемой области. Необходимо подчеркнуть, что размер прогнозируемой области должен быть больше размера исходного объекта с предыдущего кадра. Высоту объекта обозначим H , ширину — W . Размер прогнозируемой области определяется путем увеличения высоты исходного объекта в N раз и увеличения ширины объекта в R раз. Из этого следует, что размер области определяется следующим образом: $N * H$ и $R * W$, где N и R могут принимать значения от 2 до 4.

Результаты проведенных исследований приведены в табл. 4. Следует обратить внимание на то, что в зависимости от полученных результатов анализа каждого кадра, ячейки таблицы заполняются определенными символами:

- 1) полное попадание человека в прогнозируемую область обозначается (+);
- 2) частичное попадание человека обозначается (+/-);
- 3) отсутствие человека в прогнозируемой области обозначается (-).

Как демонстрирует табл. 4, при значении ширины прогнозируемой области $2W$ независимо от высоты области, объект не был обнаружен. При значении ширины прогнозируемой области $3W$ независимо от значения высоты области, алгоритм, начиная с 7 кадра, не зафиксировал объект в прогнозируемой области. Отсюда следует, что значения ($2H$ и $3W$), ($3H$ и $3W$) и ($4H$ и $3W$) оказались неэффективными для задач обнаружения объекта.

При значениях прогнозируемой области ($2H$ и $4W$), ($3H$ и $4W$) и ($4H$ и $4W$) результаты оказались успешными, на каждом кадре объект был обнаружен.

Таблица 2. Сравнительный анализ при различных значениях Дельты E
[Table 2. Comparative analysis for different values of Delta E]

Значение ΔE	№ исследуемого кадра											
	1	2	3	4	5	6	7	8	9	10	11	12
$\Delta E < 2$	-	+	-	-	-	-	-	+	+	-	-	-
$\Delta E < 5$	+	+	+	+	+	-	-	+	+	-	+	+
$\Delta E < 7$	+	+	+	+	+	+	+	+	+	+	+	+
$\Delta E < 12$	+	+	+	+	+	+	+	+	+	+	+	+
$\Delta E < 15$	-	-	+	-	-	-	+	+	+	+	-	-
Наличие объекта	Да	Да	Да	Да	Да	Да	Да	Нет	Нет	Да	Да	Да

Таблица 3. Сравнительный анализ производительности при значениях Дельты E
[Table 3. Comparative analysis of performance with selected Delta E values]

Значение ΔE	№ исследуемого кадра (время в сек)												Среднее время
	1	2	3	4	5	6	7	8	9	10	11	12	
$\Delta E < 7$	2,1	2,3	2,1	2	2	1,3	1,2	0,5	0,7	1,1	1,9	2	1,6
$\Delta E < 12$	2,6	2,8	3,1	2	2,5	2,1	1,5	0,8	0,8	1,9	2,3	2,1	2

Таблица 4. Сравнительный анализ по нахождению объекта в прогнозируемой области
[Table 4. Comparative analysis of finding an object in the predicted area]

Размер прог. обл.	№ исследуемого кадра											
	1	2	3	4	5	6	7	8	9	10	11	12
2H и 2W	-	-	-	-	-	-	-	-	-	-	-	-
2H и 3W	+	+	+	+	+/-	+/-	-	-	-	-	-	-
2H и 4W	+	+	+	+	+	+	+/-	-	-	+/-	+	+
3H и 2W	-	-	-	-	-	-	-	-	-	-	-	-
3H и 3W	+	+	+	+	+/-	+/-	-	-	-	-	-	-
3H и 4W	+	+	+	+	+	+	+/-	-	-	+/-	+	+
4H и 2W	-	-	-	-	-	-	-	-	-	-	-	-
4H и 3W	+	+	+	+	+/-	+/-	-	-	-	-	-	-
4H и 4W	+	+	+	+	+	+	+/-	-	-	+/-	+	+
Наличие объекта	Да	Да	Да	Да	Да	Да	Да	Нет	Нет	Да	Да	Да

Для выбора наиболее оптимального размера прогнозируемой области необходимо проанализировать производительность работы алгоритма при анализе каждого доступного видеокadra. Следует отметить, что, чем больше прогнозируемая область, тем больше необходимо времени для анализа цветowych пикселей в рамках данной области. Пользуясь данными табл. 5, можно сделать вывод, что при значениях размера прогнозируемой области (3H и 4W) и (4H и 4W) алгоритму

требуется больше времени для анализа, чем при значении 2H и 4W.

Таким образом, наиболее оптимальным размером прогнозируемой области является значение 2H и 4W.

Проведенные исследования позволяют сделать следующий вывод:

1) оптимальным значением Дельты E является $\Delta E < 7$;

2) оптимальный размер прогнозируемой области составляет 2H и 4W.

Таблица 5. Сравнительный анализ производительности при различных значениях размера области

[Table 5. Comparative analysis of performance at different values of the size of the area]

Размер прог. обл.	№ исследуемого кадра (время в сек)												Среднее время
	1	2	3	4	5	6	7	8	9	10	11	12	
2H и 4W	2,2	2,2	2,1	2,1	2	1,3	1,2	0,5	0,7	1,2	1,9	2	1,6
3H и 4W	2,7	2,6	2,8	2,6	2,6	1,9	1,8	0,9	0,9	2	2,7	2,8	2,1
4H и 4W	3	3,9	3,2	2,9	2,6	2,7	1,8	1	1,1	2,4	2,8	3,1	2,5

На основе полученных экспериментальных данных было проведено полное тестирование предлагаемого алгоритма. Для тестирования было выбрано видео, в котором два человека, расположенные недалеко друг от друга, перемещаются, покидают поле зрения камеры и возвращаются обратно.

После анализа каждого кадра видео программа выводит на экране первого кадра все центральные точки локализации конкретного человека с каждого кадра.

На рис. 6 отображается общая трассировка человека в красной рубашке по центру объекта, которая наглядно демонстрирует его передвижение в ходе всего видеопотока. Необходимо отметить, что в верхнем текстовом поле левой панели указываются все кадры, на которых был обнаружен потенциальный объект и координаты его расположения. А в следующем поле записываются номера кадров, в которых человек отсутствует.

По итогам исследования и тестирования был сделан вывод, что предлагаемый алгоритм работает стабильно, даже если отслеживаемый человек покидает поле зрения камеры и возвращается обратно. Это доказывает эффективность анализа цветовых значений пикселей для задач поиска и распознавания объектов. Настоящая работа является началом исследований в области трассировки объектов в видеопотоке с камер видеонаблюдения. В дальнейшем планируется теоретическое и экспериментальное исследование в области нейронных сетей. Весьма перспективным направлением дальнейшей работы является разработка алгоритмов на основе комбинаций нейронных сетей и предлагаемого алгоритма. Вторым возможным направлением может быть разработка новых алгорит-

мов, которые могут более точно и детально анализировать исследуемые объекты и прогнозировать направления их движения.

ЗАКЛЮЧЕНИЕ

Представленная статья посвящена изучению проблемы сопровождения объектов в видеопотоке. Теоретический анализ работ современных исследователей (Т. Анштедт, П. Виола, Е. Ю. Минаев, В. В. Кутикова, А. А. Артемов, С. Д. Шелабин, А. В. Пчелкина и др.) показывает, что данная проблема рассматривалась достаточно широко. Усиление внимания к данной проблеме связано с тем, что существующие в настоящее время методы обладают целым рядом недостатков, устранение которых является весьма перспективным.

В результате изучения данной проблемы был разработан алгоритм трассировки передвижений человека в видеопотоке с возможностью отображения траекторий движения. Основой предложенного алгоритма является цифровая обработка исследуемого изображения человека, состоящая из двух последовательных частей — разделения тела человека на составляющие части и получения цветовой гистограммы этих частей. Разработанный алгоритм с помощью цифровой информации объекта на каждом последующем кадре анализирует все объекты в прогнозируемой области и выполняет проверку принадлежности обнаруженных объектов к заранее выбранному объекту. Выходными данными предложенного алгоритма являются координаты локализации человека на каждом кадре и объединение этих данных для формирования и отображения общей картины трассировки исследуемого объекта.



Рис.6. Отображение трассировки объекта по центру объекта
 [Fig. 6. Displaying the object trace in the center of the object]

В итоге хотелось бы подчеркнуть следующее: предложенный алгоритм полностью автоматизирован и обладает такими преимуществами, как отсутствие необходимости в предварительном обучении, не требует использования специального оборудования, обладает способностью повторного поиска человека при его потере (перекрытие человека другими объектами и выход за пределы поля зрения камер), что позволяет использовать его для решения многих задач в области видеоаналитики.

КОНФЛИКТ ИНТЕРЕСОВ

Авторы декларируют отсутствие явных и потенциальных конфликтов интересов, связанных с публикацией настоящей статьи.

СПИСОК ЛИТЕРАТУРЫ

1. *Аништедт Т.* Видеоаналитика: мифы и реальность / Т. Аништедт, И. Келлер, Х. Лутц. – 2-е изд., перераб. и доп. – Москва : Security Focus, 2022. – 186 с.
2. *Viola P.* Rapid Object Detection using a Boosted Cascade of Simple Features / P. Viola, M. Jones // Accepted Conference on Computer Vision and Pattern Recognition (CVPR 2001). – 2001. – С. 511–518. DOI: 10.1109/CVPR.2001.990517.
3. *Минаев Е. Ю.* Трекинг объектов в видеопотоке на основе сверточных нейронных сетей и фрактального анализа / Е. Ю. Минаев,

В. В. Кутикова, А. В. Никоноров // Новая техника. – 2018. – С. 2792–2898.

4. *Артемов А. А.* Проблема поиска объектов на изображениях с помощью компьютерного зрения на основе информации о цвете / А. А. Артемов, М. В. Кавалеров, Г. С. Кузнецов // Вестник пермского государственного технического университета. Электротехника, информационные технологии, системы управления. – 2011. – С. 70–79.

5. *Шелабин Д. А.* Процесс трассировки движущихся объектов в видеопотоке / Д. А. Шелабин, С. Л. Сергеев // Актуальные проблемы гуманитарных и естественных наук. – 2013. – С. 71–77.

6. *Живрин Я. Э.* Методы определения объектов на изображениях / Я. Э. Живрин, Нафе Башар Алкзир // Молодой ученый. – 2018. – № 7 (193). – С. 8–19.

7. *Пчелкин А. В.* Алгоритм измерения координат объекта в видеопотоке для системы сопровождения объектов с прогнозом траектории движения / А. В. Пчелкин, М. Б. Никифоров // Известия Тульского государственного университета. Технические науки. – 2020. – № 12. – С. 29–35.

8. *Grabner H.* Real-time Tracking via Online Boosting / H. Grabner, M. Grabner, H. Bischof // In Proceedings British Machine Vision Conference (BMVC). – 2006. – С. 47–56. – DOI:10.5244/C.20.6.

9. *Babenko B.* Visual Tracking with Online Multiple Instance Learning / B. Babenko // Com-

puter Vision and Pattern Recognition, CVPR 2009. IEEE Conference. – 2009. – С. 101–108. – DOI: 10.1109/CVPR.2009.5206737.

10. *Henriques J. F.* High-Speed Tracking with Kernelized Correlation Filters / J. F. Henriques // IEEE transactions on pattern analysis and machine intelligence. – 2015. – С. 301–314. DOI: 10.1109/TPAMI.2014.2345390.

11. Нейросетевая система отслеживания и распознавания объектов в видеопотоке / М. В. Ляшов [и др.] // Современные наукоемкие технологии. – 2018. – № 12–1. – С. 102–107.

12. *Пастушков А. В.* Метод и алгоритмы поиска объекта в видеопотоке: автореф... дис. канд. техн. наук. – Томск : ТУСУР РАН, 2018. – 135 с.

13. *Гуменюк М. М.* Анализ решений по обнаружению и распознаванию объектов на изображениях / М. М. Гуменюк // Проблемы управления в социально-экономических и технических системах: сб. науч. статей XVIII Междунар. науч.-практ. конф. (Саратов, 14–15 апреля 2022 г.). – Саратов, 2022. – С. 126–132.

14. *Ерохин Д. Ю.* Современные сверточные нейронные сети для обнаружения и распознавания объектов / Д. Ю. Ерохин, М. Д. Ершов // Цифровая обработка сигналов. – 2018. – № 3. – С. 64–69.

15. *Кухарев Г. А.* Методы двумерной проекции цифровых и изображений в собственные подпространства: особенности реализации и

применение / Г. А. Кухарев, Н. Л. Щеголова // Компьютерная оптика. – 2018. – Т. 42, № 4. – С. 637–656.

16. *Oliva D.* Metaheuristic Algorithms for Image Segmentation: Theory and Applications / D. Oliva, M. Abd Elaziz, S. Hinojosa. – New York: Springer Publishing Company, 2019. – 226 с. https://doi.org/10.1007/978-3-030-12931-6_14.

17. Алгоритм Grabcut OpenCV: сегментация и извлечение переднего плана. – Режим доступа: <https://russianblogs.com/article/44813115929/>. – (Дата обращения: 31.01.23).

18. *Fukunaga K.* The estimation of the gradient of a density function, with applications in pattern recognition / K. Fukunaga, L. D. Hostetler // IEEE Transactions of Information Theory. – 1975. – Vol. 21. – С. 32–40. – DOI: 10.1109/TIT.1975.1055330.

19. *Шапошников А. И.* Подбор вектора признаков для процедуры MeanShift / А. И. Шапошников // Доклады Томского государственного университета систем управления и радиоэлектроники. – 2021. – Т. 24, № 2. – С. 34–38.

20. Формула цветового отличия. – Режим доступа: https://ru.wikipedia.org/wiki/Формула_цветового_отличия. – (Дата обращения: 14.01.23).

Гуменюк Михаил Михайлович — аспирант кафедры «Прикладные информационные технологии» Саратовского государственного технического университета им. Гагарина Ю. А.

E-mail: GumenRoma@mail.ru

ORCID iD: <https://orcid.org/0009-0002-0144-7747>

Бровко Александр Валерьевич — д-р физ.-мат. наук, профессор кафедры «Прикладные информационные технологии» Саратовского государственного технического университета им. Гагарина Ю. А.

E-mail: brovkoav@gmail.com

ORCID iD: <https://orcid.org/0000-0002-6956-4299>

ALGORITHM FOR TRACKING THE MOVEMENTS OF A PERSON IN THE VIDEO STREAM FROM SURVEILLANCE CAMERAS

© 2023 M. M. Gumenyuk✉, A. V. Brovko

*Yuri Gagarin State Technical University of Saratov
77, Politechnicheskaya Street, 410054 Saratov, Russian Federation*

Annotation. The task of tracking the movement of objects in the video stream (tracing) is one of the most difficult tasks of modern video surveillance and video analytics systems. The tracing task is understood as the task of automated recognition and further tracking of objects throughout the sequence of frames of the video stream. The purpose of this article is not only a brief analysis of some of the existing methods and algorithms necessary for visualizing the movement of objects, but also the development of a new algorithm for tracing human movements in a video stream with the ability to display motion trajectories. The proposed approach makes it possible to eliminate some of the shortcomings of currently existing methods (presence of special equipment, the need for preliminary training, overlapping of the object with other objects, going beyond the camera's field of view). The paper describes the main advantages of this algorithm, its most important functions and capabilities. The main stages of the algorithm operation include: selecting a person in the video stream, implementing the digital processing procedure, which is based on dividing the human body into its constituent parts and obtaining a color histogram of these parts, predicting the location and recognizing the person under study on all subsequent frames of the video stream, where the RGB pixel color analysis is performed using the method of comparing the obtained data with the color histogram of the constituent parts of the main object. The output data of the proposed algorithm is used in the procedure of forming and displaying a general picture of the movement of a particular object within the entire video stream. This article may be of interest to specialists and experts who use computer vision methods in their work to automatically obtain the necessary data when analyzing video fragments.

Keywords: surveillance cameras, pattern recognition, delta E, tracing, YOLO neural network, active area, object detection, tracking, mathematical statistics, predicted area, RGB pixels.

CONFLICT OF INTEREST

The authors declare the absence of obvious and potential conflicts of interest related to the publication of this article.

REFERENCES

1. Anstedt T., Keller I. and Lutz H. (2022) Video analytics: Myths and reality. Ed. «Security Focus». (In Russian)

2. Viola P. and Jones M. (2001) Rapid Object Detection using a Boosted Cascade of Simple Features. *Accepted Conference on Computer Vision and Pattern Recognition (CVPR 2001)*. P. 511–518. doi: 10.1109/CVPR.2001.990517

3. Minaev E. Yu., Kutikova V. V. and Nikonov A. V. (2018) Tracking of objects in a video stream based on convolutional neural networks and fractal analysis. *New technique*. P. 2792–2898. (In Russian)

4. Artemov A. A., Kavalerov M. V. and Kuznetsov G. S. (2011) The problem of finding objects in images using computer vision based on color information. *Bulletin of Perm State Technical University. Electrical engineering, information technology, control systems*. P. 70–79. (In Russian)

5. Shelabin D. A. and Sergeev S. L. (2013) The process of tracing moving objects in a video stream. *Actual problems of humanities and Natural Sciences*. P. 71–77. (In Russian)

6. Zhivrin Ya. E. and Alkzir Nafeh Bashar (2018) Methods for determining objects in images. *Young Scientist Magazine*. 7 (193). P. 8–19. (In Russian)

✉ Gumenyuk Mikhail M.
e-mail: GumenRoma@mail.ru

7. Pchelkin A. V. and Nikiforov M. B. (2020) An algorithm for measuring the coordinates of an object in a video stream for an object tracking system with a prediction of the trajectory of movement. *Izvestia of Tula State University. Technical Sciences*. (12). P. 29–35. (In Russian)
8. Grabner H., Grabner M. and Bischof M. (2006) Real-time Tracking via On-line Boosting. In *Proceedings British Machine Vision Conference (BMVC)*. P. 47–56. doi: 10.5244/C.20.6
9. Babenko B. (2009) Visual Tracking with Online Multiple Instance Learning. *Computer Vision and Pattern Recognition, CVPR 2009. IEEE Conference*. P. 101–108. doi: 10.1109/CVPR.2009.5206737
10. Henriques J. F. (2015) High-Speed Tracking with Kernelized Correlation Filters. *IEEE transactions on pattern analysis and machine intelligence*. P. 301–314. doi: 10.1109/TPAMI.2014.2345390
11. Lyashov M. V., Bereza A. N. Babaev A. M. and Kotsyubinskaya S. A. (2018) Neural network system for tracking and recognizing objects in a video stream. *Modern high-tech technologies*. (12-1), P. 102–107. (In Russian)
12. Pastushkov A. V. (2018) Method and algorithms for searching for an object in a video stream. *Tomsk State University of Control Systems and Radio electronics (TUSUR) RAS*. Tomsk, Russia. (In Russian)
13. Gumenyuk M. M. (2022) Analysis of solutions for detecting and recognizing objects in images. *Management problems in socio-economic and technical systems: sat. sci. articles XVIII International Scientific and Practical Conference*. 14–15 April 2022, Saratov, Russia, P. 126–132. (In Russian)
14. Erokhin D. Yu. and Ershov M. D. (2018) Modern convolutional neural networks for object detection and recognition. *Digital signal processing*. 3. P. 64–69. (In Russian)
15. Kukharev G. A. and Shchegolova N. L. (2018) Methods of two-dimensional projection of digital and images into their own subspaces: features of implementation and application. *Computer optics*. 42(4). P. 637–656. (In Russian)
16. Oliva D., Abd Elaziz M. and Hinojosa S. (2019) *Metaheuristic Algorithms for Image Segmentation: Theory and Applications*. New York, Springer Publishing Company. https://doi.org/10.1007/978-3-030-12931-6_14.
17. RussianBlogs. (2019) OpenCV Grabcut algorithm: segmentation and foreground extraction. Available at: <https://russianblogs.com/article/44813115929/> (accessed: 31th January 2023).
18. Fukunaga K. and Hostetler L. D. (1975) The estimation of the gradient of a density function, with applications in pattern recognition. *IEEE Transactions of Information Theory*. 21. P. 32–40. doi: 10.1109/TIT.1975.1055330.
19. Shaposhnikov A. I. (2021) Feature vector selection for the MeanShift procedure. *Reports of Tomsk State University of Control Systems and Radio electronics*. 24 (2). P. 34–38. (In Russian)
20. Wikipedia (2020) Color difference. Available at: https://en.wikipedia.org/wiki/Color_difference (accessed: 14th January 2023).

Gumenyuk Mikhail M. — PhD student, Department of Applied Information Technologies, Yuri Gagarin State Technical University of Saratov.

E-mail: GumenRoma@mail.ru

ORCID iD: <https://orcid.org/0009-0002-0144-7747>

Brovko Alexander V. — Doctor of Sciences in Physics and Mathematics, Professor of the Department of Applied Information Technologies, Yuri Gagarin State Technical University of Saratov.

E-mail: brovkoav@gmail.com

ORCID iD: <https://orcid.org/0000-0002-6956-4299>