

Алгоритм трассировки перемещений человека в видеопотоке на основе технологий распознавания одежды

М.М. Гуменюк, А.В. Бровко

Саратовский государственный технический университет им. Гагарина Ю.А.

Аннотация: В настоящее время в системах видеонаблюдения и видеоаналитики центральное место занимает трассировка перемещений различных объектов (в частности, человека). Она представляет собой систему отслеживания перемещений людей путем локализации их позиций на каждом кадре в рамках всего видеопотока и является основой множества интеллектуальных систем компьютерного зрения. Цель данной статьи - разработка нового алгоритма трассировки перемещений человека в видеопотоке с возможностью отображения траекторий движения. Основные этапы работы алгоритма включают в себя: разделение видео на кадры с разницей в одну секунду, выбор исследуемого человека в видеопотоке, осуществление процесса цифровой обработки, основой которого является распознавание одежды исследуемого человека и получение ее цветовой гистограммы, прогнозирование локализации и распознавание исследуемого человека на всех последующих кадрах видеопотока с помощью разработанной методики прогнозирования направления движения данного объекта. Выходные данные предложенного алгоритма применяются в процедуре формирования и отображения общей картины перемещения конкретного человека в рамках всего видеопотока. Информация и материалы, содержащиеся в данной статье, могут представлять интерес для специалистов и экспертов, которые в своих работах особое внимание уделяют обработке данных при анализе фрагментов видеопотока.

Ключевые слова: камеры наблюдения, нейронная сеть U^2 -net, библиотека `rembg`, распознавание образов, распознавание одежды, дельта E, трассировка, прогнозирование направления движения, обнаружение объектов, трекинг, математическая статистика, прогнозируемая область, пиксели RGB.

Введение

Актуальность темы

Современный мир демонстрирует конвергенцию различных сфер деятельности человека. Эта тенденция отчетливо прослеживается в области видеонаблюдения, куда все более интенсивно проникают информационные технологии. Видеонаблюдение и охранное телевидение заимствуют различные компьютерные и сетевые технологии, на основе которых разрабатываются уже совершенно новые интеллектуальные алгоритмы, позволяющие расширить функциональность и повысить эффективность

систем видеонаблюдения. Другими словами, отрасль видеонаблюдения постоянно вбирает в себя все новейшие достижения науки и техники.

Современные системы видеонаблюдения немислимы без применения интеллектуальных технологий обработки данных, позволяющих в реальном времени анализировать не только отдельные изображения, но и целые последовательности динамических событий и сцен. В целом можно отметить, что видеоаналитика является неотъемлемой частью систем видеонаблюдения. Видеоаналитика представляет собой аппаратно-программное обеспечение или технологию, которая использует различные современные методы компьютерного зрения для автоматизированного сбора данных на основании анализа потокового видео. В сфере видеоаналитики центральное место занимают методы и алгоритмы обработки исследуемого изображения и распознавания образов, позволяющие производить процесс анализа без непосредственного участия человека.

Необходимо обратить внимание на то, что в настоящее время существует множество различных функциональных возможностей видеоаналитики, основными из которых являются:

- 1) функция улучшения изображений [1, 2];
- 2) функция детектирования движения [1, 2];
- 3) функция распознавания лиц [3];
- 4) функция классификации объектов [3];
- 5) функция трассировки объекта или группы объектов [4, 5];
- 6) функция распознавания автомобильных номеров [4, 5].

Необходимо подчеркнуть, что на основе вышеперечисленных базовых функциональных возможностей и их комбинаций могут быть созданы разнообразные методы и новые функции для анализа фрагментов видеопотока.

В текущий момент времени одной из наиболее сложных задач современных систем видеонаблюдения и видеоаналитики является трассировка (трекинг) перемещений объекта или группы объектов в видеопотоке на основе их визуальных характеристик. Под задачей трассировки понимается задача автоматизированного распознавания и дальнейшего отслеживания объектов на протяжении последовательности кадров видеопотока. Данный процесс отслеживания перемещения объектов применяется во многих задачах видеоаналитики, среди которых наиболее важными и актуальными в настоящее время являются:

- мониторинг дорожных ситуаций (например, измерение интенсивности дорожного движения);
- проведение статистических исследований в торговых и развлекательных центрах, на промышленных предприятиях;
- осуществление автоматического контроля воздушной обстановки;
- контроль доступа на охраняемые объекты;
- анализ работы сотрудников государственных и частных предприятий.

Первостепенное значение для рассмотрения данного вопроса имеет изучение работ различных современных исследователей (Т. Анштедт, А. А. Лукьяница, А. Г. Шишкин, П. Виола, А. А. Артемов, Е. Ю. Минаев, М. В. Кавалеров, С. Д. Шелабин, В. В. Кутикова, Я. Э. Живрин, А. В. Пчелкина и др.) по данной тематике. Глубокое и всестороннее рассмотрение различных аспектов подчеркивает тот факт, что в настоящее время задача трассировки (трекинга) объектов не является полностью решенной [6, 7]. Усиление внимания к проблеме сопровождения объектов в видеопотоке [8-10] связано с расширением сфер применения трекинга (трассировки), с увеличением вычислительных мощностей процессоров ЭВМ и электронно-вычислительных средств цифровой обработки изображений [11], которое дает возможность реализовывать все более сложные алгоритмы и оснащать

создаваемые устройства новыми функциями, практически граничащими с ожидаемыми возможностями искусственного интеллекта, а также с наличием определенных сложностей (изменения условий освещенности, проективные и аффинные искажения, перекрытие объектов другими объектами и другие) [12, 13]. Таким образом, вышеперечисленные факты приводят к необходимости разработки новых или модификации существующих в настоящее время методов и алгоритмов трассировки объектов в видеопотоке.

Существующие методы и их недостатки

Теоретический анализ литературы показывает, что данный вопрос рассматривался достаточно широко. Необходимо отметить, что некоторые существующие в настоящее время современные методы визуализации движения объектов в видеопотоке обладают следующими недостатками:

- методы (Online Boosting Tracker, MIL Tracker, KCF Tracker) обладают таким недостатком, как отсутствие возможности повторного поиска объекта при его потере в видеокадрах (перекрытие другими объектами или выход за рамки кадров) [14-16];
- методы (Google Class, OpenCV, Vuforia SDK, Kinect SDK) устраняют вышеуказанный недостаток, но и в свою очередь обладают другими недостатками — необходимостью наличия специального оборудования и длительностью процесса предварительного обучения [17, 18].

Предлагаемый способ решения задач

Представленная статья посвящена детальному описанию предлагаемого авторами алгоритма трассировки перемещений человека. Особенностью данного алгоритма является применение специальных технологий распознавания одежды исследуемого человека на основе нейронной сети U^2 –сети. Кроме того, для анализа видеокадров применяется методика создания прогнозируемых областей и подобластей, которые

определяются путем прогнозирования направления движения исходного объекта. Предложенный алгоритм обладает следующими преимуществами:

- 1) отсутствие необходимости предварительного обучения и использования специального оборудования;
- 2) возможность повторного поиска при потере исследуемого человека при перекрытии его другими объектами или при его выходе за пределы поля зрения камер.

Описание предлагаемого алгоритма

Предлагаемый алгоритм представляет собой трудоемкий процесс вычисления и создания трассировки исследуемого человека в ходе всего видеопотока. Данный процесс разбивается на множество последовательных этапов, основные из которых представлены в блок-схеме на рис. 1.

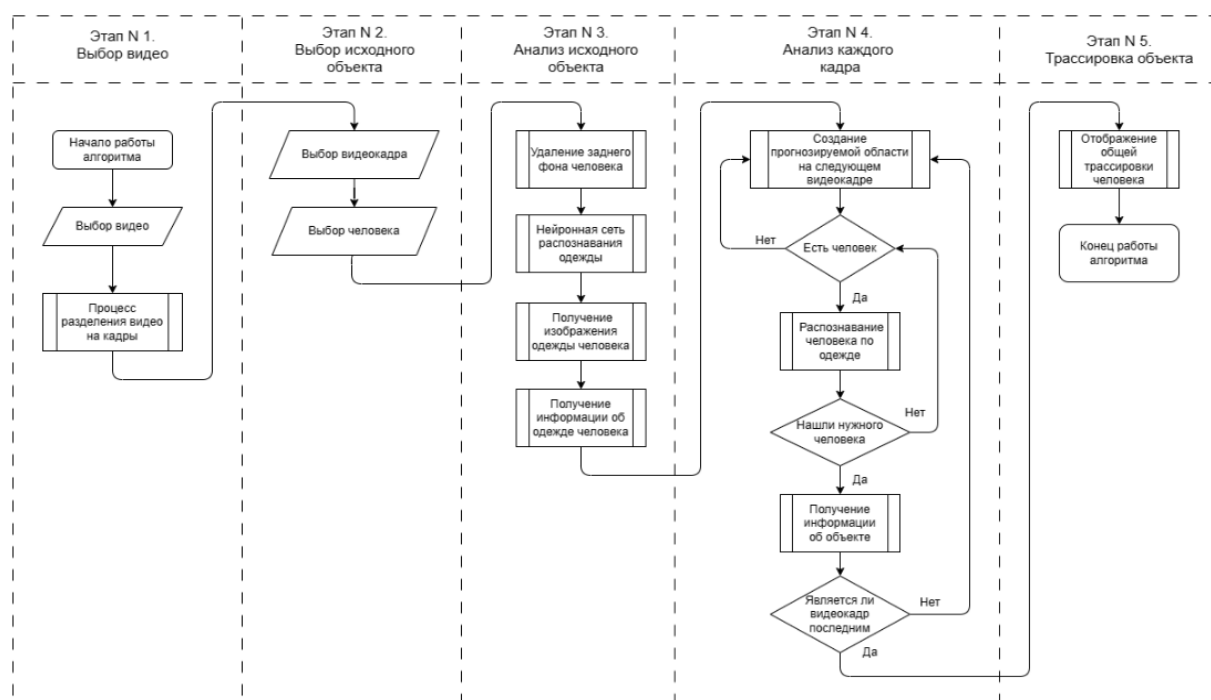


Рис. 1. – Блок-схема основных этапов предложенного алгоритма

Для разработки и тестирования предлагаемого алгоритма было разработано программное обеспечение, состоящее из множества последовательных и взаимосвязанных между собой программ,

реализованных на языке C# и Python в среде Visual Studio 2022. Необходимо уточнить, что все последующие рисунки и таблицы, представленные в данной статье, демонстрируют результаты работы программного обеспечения, тестирующего предложенный алгоритм.

Разделение видео на кадры

На данном этапе исследуемое видео анализируется по таким параметрам, как формат видео, разрешение, длина, общее количество кадров всего видеопотока и кадровая частота. Кадровая частота представляет собой количество статических изображений, сменяющихся в течение одной секунды на экране и создающих иллюзию движения в видеопотоке. С целью оптимизации алгоритма был предложен способ выборочного анализа кадров, который заключается в разделении видеопотока на множество кадров с разницей в 1 секунду. Данная стратегия была предложена для того, чтобы для слежения за объектом не было необходимости рассматривать все кадры видеопотока, а только те кадры, которые были отобраны согласно определенным требованиям. Необходимо отметить, что предложенный способ позволит ускорить процедуру анализа видеок кадров и уменьшить размеры хранения информации.

Для наглядности приведем пример видео, основными характеристиками которого являются — длительность 1 час и 15 минут, формат avi, 24 кадра в секунду. В табл. 1 указано количество кадров при полном и выборочном анализе. Данная таблица наглядно демонстрирует, что при выборочном анализе будут проанализированы всего 4 500 кадров вместо 108 000 кадров, что значительно ускорит работу по анализу видео.

Таким образом, пользователь должен загружать интересующее его видео в разработанную авторами программу. Алгоритм программы с помощью специальных методов библиотеки OpenCV на языке Python анализирует и сохраняет в созданную папку все кадры, которые соответствуют требованиям

отбора кадров. После выполнения данной операции алгоритм приступает к следующему этапу – этапу выбора исходного объекта.

Выбор исходного объекта

На этом этапе основной задачей является выбор человека, для которого необходимо построить траекторию перемещения путем локализации позиций на каждом кадре в рамках всего видеопотока. На первичном видеокадре пользователь самостоятельно с помощью специального инструмента выделяет интересующего его человека. Данная процедура необходима для того, чтобы алгоритм имел возможность проанализировать исследуемое изображение для получения конкретной информации, необходимой для дальнейшей работы предложенного алгоритма трассировки. Как только объект был выбран, разработанный алгоритм приступает к этапу его анализа.

Таблица №1

Сравнение количества кадров при полном и выборочном анализе

	Исследуемое видео			
	Количество кадров в секунду	Количество кадров в минуту	Количество кадров в час	Количество кадров за 1 час 15 минут
Полный анализ	24 кадра	1 440 кадров	86 400 кадров	108 000 кадров
Выборочный анализ	1 кадр	60 кадров	3 600 кадров	4 500 кадров

Анализ исходного объекта

Данный этап начинается с процесса выделения переднего плана. Этот процесс заключается в отделении движущихся фрагментов изображения от неподвижных, которые называют фоновыми или принадлежащими заднему плану. От того, насколько аккуратно и корректно решена эта задача, зависят все последующие этапы обработки информации, а также требуемые вычислительные ресурсы. Таким образом, этапу выделения

переднепланового изображения и применяемым для этого методам уделено особое внимание.

Следует подчеркнуть, что в алгоритме, представленном нами в статье «Алгоритм отслеживания перемещений человека в видеопотоке с камер видеонаблюдения» для отделения человека от заднего фона применялся специальный шаблон [19]. Шаблон представляет собой некий видеокادر, на котором должны отсутствовать какие-либо объекты. Данный способ состоит в сопоставлении изображения человека с шаблоном, результатом которого является отделение человека от заднего фона. Недостатком такого подхода является то, что пользователю необходимо просмотреть все видеокadres и выбрать конкретный кадр, который в дальнейшем послужит шаблоном.

Так как проблема отделения объекта (человека) от заднего фона является предметом дополнительного исследования, выходящего за рамки нашей работы, мы сконцентрируем свое внимание на рассмотрении существующих методов и алгоритмов для отделения объекта от заднего фона. Необходимо обратить внимание на то, что были проанализированы только те алгоритмы, которые являются полностью автоматизированными, иными словами, не требующими вмешательства пользователя в процесс анализа. Не вдаваясь в подробности анализа методов, отметим, что наилучший результат демонстрирует современная специальная библиотека Rembg, созданная с помощью языка программирования Python. Эта библиотека была разработана Даниэлем Гатисом в 2020 году на основе U^2 -сети [20]. U^2 -сеть представляет собой глубокую нейронную сеть для обнаружения видимых объектов (SOD). Основой ее архитектуры является двухуровневая U – образная структура, состоящая из многоуровневой глубокой интеграции функций и масштабного извлечения объектов [21].

Таким образом, данная стадия процесса анализа исходного объекта состоит в применении библиотеки Rembg, позволяющей алгоритму отделять

человека от заднего фона. Необходимо подчеркнуть, что в данном процессе применяется так называемое матирование изображения (мягкое выделение переднего фона), которое способно с точностью до пикселя удалять оставшиеся области заднего фона на переднем фоне человека. Данный подход необходим для более точного распознавания одежды на исходном объекте.

Следующим этапом процесса анализа исходного (оригинального) объекта является обнаружение и выделение одежды объекта. Здесь применяется библиотека Cloth Segmentation, разработанная на основе U^2 -сети. Данная библиотека была создана в 2021 году Левином Дахби и бесплатно доступна в GitHub [22]. Она включает в себя обучающий код, код вывода и предварительно обученную модель для синтаксического анализа тканей по изображению человека. Необходимо отметить, что данная модель выводит четыре канала, каждый из которых представляет общий фон, всю одежду человека, верхнюю и нижнюю части одежды. Кроме того, данная библиотека не сможет правильно распознавать одежду, если на исследуемом изображении присутствуют больше одного человека.

Нейронная сеть сегментации одежды анализирует входящее изображение, в данном случае человека с прозрачным фоном. В случае успешного обнаружения одежды человека, данная обученная модель создает особое изображение, которое называется маской изображения. Это такое изображение, интенсивность пикселей которого равна либо нулю, либо единице, другими словами, двоичная интенсивность. Маска представляет собой изображение, состоящее только из двух цветов — черного (интенсивность пикселей 0) и белого (интенсивность пикселей 1).

Для получения цветовой информации об одежде необходимо рассматривать только одежду исследуемого человека и исключать все остальные части тела человека. Для этого применяется маска, которая

накладывается на первоначальное изображение. Данный процесс состоит в следующем — цветовые пиксели оригинального изображения человека сверяются с пикселями маски. Если пиксели маски соответствуют значению 1, то пиксели оригинального изображения не подвергаются изменениям, при значении 0 пиксели оригинального изображения игнорируются. Таким образом, благодаря маске создается новое изображение одежды и алгоритм приступает к заключительному этапу анализа исходного объекта — этапу получения цветовой информации об одежде исследуемого человека.

Следует отметить, что для получения цветовой информации об одежде на данном этапе предложенный алгоритм применяет один из самых популярных на сегодняшний день цветовых форматов — RGBA. RGBA расшифровывается как красный-зеленый-синий-альфа и представляет собой цветовое пространство, состоящее из адаптивной трехканальной цветовой модели RGB и четвертого альфа-канала. RGB был разработан как модель получения любых оттенков видимого человеческим глазом спектра, в которой смешиваются красный (Red), зеленый (Green) и синий (Blue) каналы с разной интенсивностью излучения. А альфа-канал позволяет определять степень прозрачности цветового пикселя. Подчеркнем, что для получения изображения одежды исследуемого человека задний фон заменяется на прозрачные пиксели, поэтому для анализа данного изображения был выбран формат RGBA. Благодаря этому формату, можно отделить все прозрачные пиксели от оставшихся цветовых пикселей изображения одежды исследуемого человека. Еще одной важной чертой является использование формата файлов растровых изображений. Рекомендуется все изображения преобразовывать в формат PNG, поскольку данный формат поддерживает цветовое пространство RGBA, в отличие от JPG, в котором прозрачные пиксели заменяются на пиксели черного цвета.

После получения всех непрозрачных цветовых пикселей изображения одежды исследуемого человека можно перейти от RGBA к цветовому пространству RGB, поскольку альфа-каналы в дальнейшем не будут учитываться.

Прежде чем анализировать и отбирать пиксели одежды объекта, необходимо учитывать два факта:

1) пиксели одной и той же одежды исследуемого человека на последующих кадрах могут различаться из-за смещения оттенков цветов и изменения уровня интенсивности освещения;

2) одежда может содержать в себе десятки тысяч разных значений цветов.

Отсюда следует, что на данном этапе необходимо подобрать определенные диапазоны цветов, другими словами, все близкие по значению цвета включить в определенную группу цветов. Этот способ необходим для того, чтобы при поиске объекта на последующих кадрах не нужно было рассматривать все цветовые пиксели одежды исходного объекта, а только те группы цветов, которые были обнаружены во время анализа исходного объекта. Для этого было предложено использовать формулу цветового отличия для формата RGB, позволяющую вычислить различие между двумя цветовыми пикселями. Такое различие между двумя цветами называют Дельтой E (Delta E) и соответствует шкале от 0 до 100. Необходимо отметить, что чем ниже показатель Дельты E, тем выше схожесть двух цветов. Дельта E вычисляется по формуле:

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

Во время анализа одежды вначале производится запись всех существующих цветовых пикселей исследуемого изображения кроме прозрачных пикселей, после чего отбираются все уникальные и неповторяющиеся пиксели. Далее с помощью формулы цветового отличия

создаются группы уникальных цветов одежды. Процесс создания особых групп цветов заключается в том, что в списке уникальных цветов пикселей изображения одежды исследуемого человека сравниваются значения всех цветовых пикселей. Если различие между ними соответствует $0 \leq \Delta E \leq 7$, то такие значения считаются одним цветом и записываются в отдельную группу цветов. Процесс продолжается до тех пор, пока не будут созданы различные группы цветов и просмотрены все значения цветовых пикселей. Полученная цветовая информация об одежде исследуемого человека позволит находить и распознавать его на последующих кадрах видеопотока.

Таким образом, последовательность основных этапов анализа исходного объекта можно представить следующим образом:

- 1) применение нейронной сети для удаления заднего фона объекта;
- 2) использование нейронной сети для распознавания одежды;
- 3) создание маски одежды;
- 4) наложение маски на изображение исходного объекта, результатом которого является получение изображения одежды исходного человека;
- 5) замена черного фона объекта на прозрачный фон;
- 6) получение цветовой информации об одежде исследуемого человека с помощью формулы цветового отличия.

На рис. 2 представлена последовательность всех основных этапов процесса анализа исходного объекта.

Анализ каждого кадра

Следующий этап является наиболее сложной частью всего алгоритма и разделяется на множество подзадач, основными из которых являются:

- 1) создание прогнозируемой области на текущем видеокadre;
 - 2) удаление заднего фона в прогнозируемой области;
 - 3) создание прогнозируемой подобласти;
-

- 4) анализ и распознавание исследуемого человека в прогнозируемой области или подобласти;
- 5) получение координат локализации обнаруженного человека и анализ следующего видеокadra.



Рис. 2. – Основные этапы анализа исходного объекта

Прогнозируемая область представляет собой некую область видеокadra, в которой может находиться рассматриваемый нами человек. Применение прогнозируемой области позволяет алгоритму осуществлять поиск исследуемого человека строго в прогнозируемой области, а не по всему кадру, размер которого может значительно превышать размер области. Необходимо подчеркнуть, что центром такой прогнозируемой области является центр локализации обнаруженного объекта с предыдущего кадра. Кроме того, размер объекта с предыдущего видеокadra определяет размер всей прогнозируемой области на текущем кадре.

Для определения наиболее оптимального размера прогнозируемой области были проанализированы данные исследований, изложенных в статье «Алгоритм отслеживания перемещений человека в видеопотоке с камер видеонаблюдения» [19]. Для выбора оптимального размера прогнозируемой области рассматривается объект, выделенный на предыдущем кадре. Высоту объекта обозначим H , ширину объекта — W . Размер прогнозируемой

области определяется путем увеличения высоты объекта в N раз и увеличения ширины объекта в R раз. Таким образом, размер данной области определяется следующим образом: $N*H$ и $R*W$, где N и R могут принимать значения от 2 до 4. Основываясь на экспериментальных данных, представленных в вышеуказанной статье, можно сделать вывод, что наиболее оптимальным размером прогнозируемой области является значение $2H$ и $4W$.

После создания прогнозируемой области на текущем кадре разработанный авторами алгоритм приступает к следующему этапу, в рамках которого был предложен дополнительный способ, позволяющий производить более быстрый поиск исследуемого объекта. Введем такое понятие, как прогнозируемая подобласть. Это такая область в прогнозируемой области, которая определяется с помощью прогнозирования движения рассматриваемого человека. Прогнозирование движения определяется путем сопоставления координат местоположения исследуемого объекта с двух последних предыдущих кадров.

Рассмотрим, на каких видеокадрах и при каких условиях применяется прогнозируемая область и подобласть:

- на первом кадре пользователь выбирает интересующий его объект;
 - во время процесса анализа второго видеокадра рассматривается вся прогнозируемая область, поскольку на данном этапе доподлинно неизвестно, в каком направлении движется рассматриваемый человек;
 - начиная с третьего видеокадра, когда известно местоположение объектов с предыдущих кадров, алгоритм будет анализировать в первую очередь только прогнозируемую подобласть. В случае успеха, алгоритм приступает к следующему кадру видеопотока. Если в прогнозируемой подобласти исследуемый человек не был обнаружен, то в этом случае алгоритм анализирует оставшуюся часть прогнозируемой области;
-

- если в прогнозируемой области исследуемый человек не был обнаружен, то размер прогнозируемой области сохраняется на следующем кадре. Предполагается, что объект был перекрыт другими объектами.

Хотелось бы подчеркнуть, что изложенный выше подход позволяет затратить гораздо меньше ресурсов и времени на обнаружение объекта по сравнению с процессом анализа последующих кадров, где на каждом видеокadre анализируется целиком вся прогнозируемая область.

Рассмотрим процесс получения результатов прогнозирования направления движения исследуемого объекта, который заключается в определении угла направления.

На данном этапе применяются координаты центров локализации объекта с двух предыдущих кадров – x_0, y_0 и x_1, y_1 . Пусть x_0, y_0 будут центром окружности двумерного пространства и началом вектора, а x_1, y_1 представляют собой координаты конца вектора.

Для определения угла между вектором и осью применяется формула:

$$\alpha = \operatorname{arctg}\left(\frac{y}{x}\right)$$

Пусть $y = y_1 - y_0$ и $x = x_1 - x_0$, тогда угол α можно определить следующим образом:

$$\alpha = \begin{cases} \operatorname{arctg}\left(\frac{y}{x}\right), & x > 0, y \geq 0 \\ \operatorname{arctg}\left(\frac{y}{x}\right) + 2\pi, & x > 0, y < 0 \\ \operatorname{arctg}\left(\frac{y}{x}\right) + \pi, & x < 0 \\ \frac{\pi}{2}, & x = 0, y > 0 \\ \frac{3\pi}{2}, & x = 0, y < 0 \\ - & x = 0, y = 0 \end{cases}$$

Необходимо отметить, что, если угол α не определен, то исследуемый объект остался на месте.

На рис. 3 представлена схема прогнозируемой области и подобластей в зависимости от результатов прогнозирования направления движения исследуемого объекта.

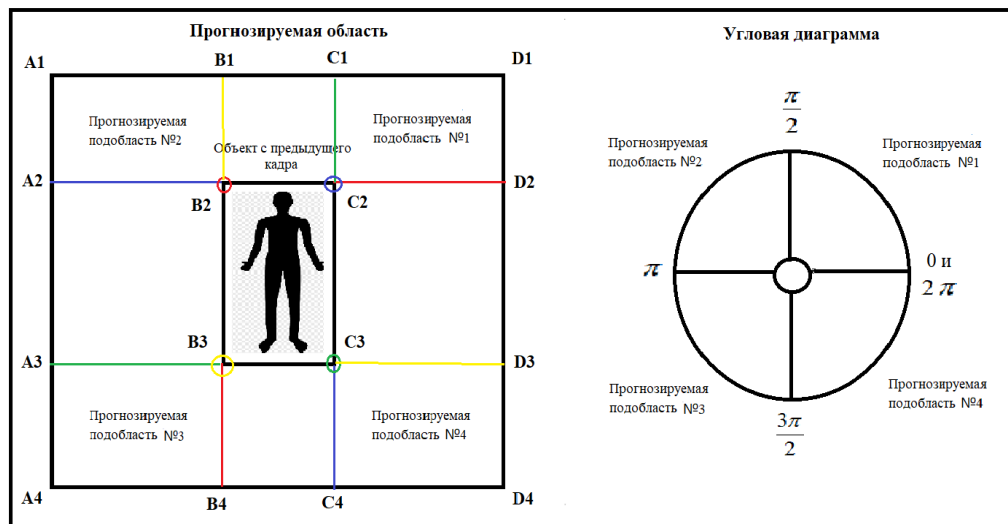


Рис. 3. – Схема получения прогнозируемой подобласти

Рассмотрим данную схему, которая наглядно показывает, при каких значениях угла α алгоритм отбирает одну из четырех возможных прогнозируемых подобластей для дальнейшего поиска исследуемого человека:

- 1) при значениях $0 \leq \alpha \leq \frac{\pi}{2}$ исследуется прогнозируемая подобласть № 1 (B1,D1,D3,B3);
- 2) при значениях $\frac{\pi}{2} < \alpha \leq \pi$ отбирается прогнозируемая подобласть № 2 (A1,C1,C3,A3);
- 3) при значениях $\pi < \alpha \leq \frac{3\pi}{2}$ анализируется прогнозируемая подобласть № 3 (A2,C2,C4,A4);

4) при значениях $\frac{3\pi}{2} < \alpha < 2\pi$ рассматривается прогнозируемая подобласть № 4 (B2,D2,D4,B4).

После получения прогнозируемой области и конкретной подобласти, предложенный авторами алгоритм приступает к следующему этапу – удалению заднего фона прогнозируемой области или подобласти. Для этого на данном этапе применяется ранее описанная в разделе данной статьи «Анализ исходного объекта» современная специальная библиотека Rembg. Основным преимуществом применения данной библиотеки является выделение всех людей в кадре и удаление заднего фона.

Таким образом, для нахождения и распознавания одежды человека алгоритм будет анализировать все цветовые пиксели людей, кроме прозрачных пикселей заднего фона. Подобная процедура позволит значительно ускорить и упростить процесс анализа по сравнению с анализом пикселей всей прогнозируемой области в целом.

После удаления заднего фона в прогнозируемой области и определения прогнозируемой подобласти, предложенный авторами алгоритм приступает к последнему этапу анализа видеокadra — этапу поиска, распознавания и определения локализации исследуемого объекта на текущем кадре.

На этом этапе применяется цветовая информация, полученная при анализе изображения одежды исходного объекта. Производится сравнение групп цветных пикселей прогнозируемой подобласти с группами цветов одежды исходного человека с помощью формулы цветового отличия. Результатом обнаружения и распознавания одежды человека является получение его начальных координат X и Y , длины и высоты. Эти данные в дальнейшем будут применяться в процедуре создания и отображения общей картины передвижения выбранного человека в рамках всего видеопотока.

Далее рассмотрим последовательность действий алгоритма в случае потери исследуемого человека. Если исследуемый объект в прогнозируемой подобласти не был обнаружен, то алгоритм приступает к анализу оставшейся части прогнозируемой области.

Если в прогнозируемой области не был обнаружен рассматриваемый объект, то предложенный алгоритм предполагает, что объект либо вышел за пределы кадра, либо был перекрыт другими объектами. Дальнейший анализ зависит от расположения прогнозируемой области на видеокадре:

1) если область находится на границе кадра — следовательно, объект вышел за пределы кадра. Не следует исключать вероятности того, что 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);

2) если область находится не на границах кадра — в таком случае, алгоритм предполагает, что объект был перекрыт другими объектами. В этом случае на следующем кадре прогнозируемая область сохраняет свой размер.

Введем следующие обозначения. Видеокадр обозначим $ABCD$, а прогнозируемую область - $A_1B_1C_1D_1$, тогда:

$$\begin{cases} A_1B_1C_1D_1 \in ABCD, & \text{— объект перекрыт} \\ A_1B_1C_1D_1 \notin ABCD, & \text{— объект вышел} \end{cases}$$

После анализа всех кадров видеопотока алгоритм предоставляет пользователю возможность построить общую трассировку перемещений исследуемого человека. Для этого алгоритм выводит на первом видеокадре все центральные точки локализации исследуемого объекта, которые были получены в процессе анализа каждого кадра. Полученный кадр наглядно демонстрирует передвижение исследуемого человека в рамках всего

видеопотока.

Результаты и обсуждение

В данном разделе рассмотрим функциональные возможности, основные преимущества предлагаемого авторами алгоритма трассировки перемещений человека на основе технологий распознавания одежды.

В таблицах № 2 и № 3 представлена сравнительная характеристика разработанных авторами алгоритмов.

Таблица № 2

Преимущества алгоритма данной статьи по сравнению с алгоритмом №1

Функции алгоритмов	Наименование алгоритма	
	Алгоритм на основе технологий распознавания одежды (Алгоритм № 2)	Алгоритм на основе разделения изображения человека на три составляющие части (Алгоритм № 1)
Прогнозирование движения объекта	Да	Нет
Автоматическое удаление заднего фона	Да	Нет
Анализ всей прогнозируемой области	При необходимости да	Да
Анализ прогнозируемой подобласти	Да	Нет
Анализ заднего фона	Нет	Да

Таблица № 3

Сравнение производительности двух разработанных авторами алгоритмов

Наименование алгоритма	№ исследуемого видеокadra												Среднее время (в сек)
	1	2	3	4	5	6	7	8	9	10	11	12	

Алгоритм №1	2, 1	2, 2	2, 3	2, 3	2	1, 3	1, 3	0,7	0,6	1, 4	1, 9	2, 1	1,6 сек
Алгоритм №2	2, 4	1, 5	1, 6	1, 7	1, 7	0, 8	0, 7	0,3	0,3	1	1, 6	1, 8	1,3 сек
Наличие объекта	да	да	да	да	да	да	да	нет	нет	да	да	да	

- Алгоритм трассировки перемещений человека на основе разделения изображения человека на три составляющие части, представленный в статье [19], обозначим, как алгоритм № 1;
- алгоритм трассировки перемещений человека на основе технологий распознавания одежды будем рассматривать как алгоритм № 2.

Алгоритм № 2 обладает рядом преимуществ по сравнению с алгоритмом № 1. Все основные преимущества продемонстрированы в табл. 2.

Следующим этапом исследования является сравнение двух представленных алгоритмов с точки зрения производительности. В таблице № 3 продемонстрированы данные, полученные нами в ходе исследования. Согласно данным, представленным в таблице № 3, среднее время анализа каждого кадра предлагаемого алгоритма № 2 составляет 1,3 секунды, в то время как среднее время работы алгоритма № 1 составляет 1,6 секунды. Таким образом, можно сделать вывод, что удаление заднего фона и создание прогнозируемой подобласти с помощью прогнозирования движения на основе данных об объекте с двух предыдущих видеок кадров являются ключевыми и дают возможность алгоритму № 2 производить поиск исследуемого человека на видеок кадрах значительно быстрее, чем алгоритм № 1.

По итогам исследования и тестирования нами был сделан вывод о том, что предлагаемый алгоритм на основе технологий распознавания одежды стабильно работает в случае перекрытия исследуемого человека другими

объектами или при выходе за пределы поля зрения камер. Это доказывает эффективность анализа цветовых значений пикселей для задач поиска и распознавания объектов. Кроме того, данный алгоритм демонстрирует более высокую скорость работы по прогнозированию и локализации объектов по сравнению с алгоритмом, представленным в статье [19].

Настоящая работа является продолжением исследований в области разработки методов и алгоритмов трассировки объектов в видеопотоке с камер видеонаблюдения.

В дальнейшем планируется теоретические и экспериментальное исследование в области нейронных сетей и цветовой гистограммы (с применением различных цветовых пространств, таких как RGB, RGBA, HSV и других) исследуемых объектов. Например, весьма актуальным и перспективным является изучение вопроса касающегося анализа частоты пикселей. Под этим понятием подразумевается то, как часто и какие цветовые пиксели в выбранных областях следует рассматривать в первую очередь (каждый пиксель, каждые два пикселя, каждые пять пикселей и далее). Вышеизложенное подчеркивает необходимость обстоятельного теоретического и практического изучения вопроса об анализе и разработке таких способов, которые позволят ускорить процесс анализа видеокадров, но и при этом не потерять в точности обнаружения исследуемых объектов. Кроме того, заслуживающим внимания направлением работы является разработка новых методов и алгоритмов на основе комбинаций нейронных сетей и разработанных авторами алгоритмов.

Заключение

Представленная статья посвящена изучению проблемы сопровождения объектов в видеопотоке и детальному описанию нового предлагаемого авторами алгоритма трассировки перемещений человека. Для более полной характеристики рассматриваемой проблемы был проведен теоретический

анализ работ современных специалистов и исследователей, таких как — Т. Анштедт, А. А. Лукьяница, А. Г. Шишкин, П. Виола, А. А. Артемов, Е. Ю. Минаев, М. В. Кавалеров, С. Д. Шелабин, В. В. Кутикова, Я. Э. Живрин, А. В. Пчелкина и др. Работы вышеперечисленных исследователей внесли серьезный вклад в развитие видеоаналитики, однако по-прежнему актуальной является проблема слежения за объектом или группой объектов в видеопотоке с камер видеонаблюдения. Усиление внимания к данной проблеме связано прежде всего с тем, что существующие в настоящее время методы и алгоритмы обладают целым рядом недостатков, устранение которых является весьма перспективным.

В результате тщательного изучения данной проблемы был разработан новый алгоритм трассировки передвижений человека в видеопотоке с возможностью отображения его общей траектории движения. Ключевой частью предложенного алгоритма является полная цифровая обработка исследуемого изображения человека, состоящая из двух последовательных этапов — этапа применения технологий распознавания и извлечения одежды исследуемого человека и этапа получения ее цветовой гистограммы. Разработанный алгоритм на каждом последующем кадре на основании данных, полученных из цифровой обработки исходного объекта, анализирует все объекты в прогнозируемой области и подобласти, в которых производится проверка принадлежности обнаруженных объектов к исследуемому исходному объекту. Кроме того, для сужения областей поиска объекта используется методика прогнозирования направления движения исследуемого человека. Выходными данными предложенного алгоритма являются координаты местоположения выбранного человека на каждом видеокadre и объединение полученных данных с целью формирования и отображения общей картины трассировки исследуемого объекта.

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). Представленный авторами материал может быть полезен для решения многих задач в области видеоаналитики.

Литература

1. Архипова Е.А., Поляничко К.С. Тенденции развития видеоаналитики в мире // Томский государственный университет систем управления и радиоэлектроники. 2020. № 7-2. С. 69-72.

2. Никитин Д.В., Тараненко И.С., Катаев А.В. Детектирование дорожных знаков на основе нейросетевой модели YOLO // Инженерный вестник Дона, 2023, №8. URL: ivdon.ru/ru/magazine/archive/n7y2023/8531.

3. Могилин К.А., Селищев В.А. Интеллектуальные системы видеонаблюдения в комплексах безопасности // Известия Тульского государственного университета. Технические науки. 2020. №3. С. 89-93.

4. Мосева М.С., Харрасов К.Р. О существующих методах удаления зашумлений на изображении // Инженерный вестник Дона, 2023, №8. URL: ivdon.ru/ru/magazine/archive/n8y2023/8643.

5. Гуменюк М.М. Анализ решений по обнаружению и распознаванию объектов на изображениях // 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.

6. Анштедт Т., Келлер И., Лутц Х. Видеоаналитика: мифы и реальность 2-е изд., перераб. и доп. Москва: Security Focus. 2022. 186 с.

7. Лукьяница А.А., Шишкин А.Г. Цифровая обработка видеоизображения. Москва: Ай-Эс-Эс Пресс. 2009. 518 с.

8. Минаев Е.Ю., Кутикова В.В., Никоноров А.В. Трекинг объектов в видеопотоке на основе сверточных нейронных сетей и фрактального анализа // Сборник трудов IV международной конференции и молодежной школы «Информационные технологии и нанотехнологии»(ИТНТ-2018)-Самара: Новая техника. 2018. С. 2792-2798.

9. Viola P., Jones M. Rapid Object Detection using a Boosted Cascade of Simple Features // Accepted Conference on Computer Vision and Pattern Recognition (CVPR 2001). 2001. Т.1. С. 511-518.

10. Шелабин Д.А., Сергеев С.Л. Процесс трассировки движущихся объектов в видеопотоке // Актуальные проблемы гуманитарных и естественных наук. 2013. №7-1. С. 71-77.

11. Артемов А.А., Кавалеров М.В., Кузнецов Г.С. Проблема поиска объектов на изображениях с помощью компьютерного зрения на основе информации о цвете // Вестник пермского государственного технического университета. Электротехника, информационные технологии, системы управления. 2011. №5. С. 70-79.

12. Живрин Я.Э., Алкзир Н.Б. Методы определения объектов на изображениях // Молодой ученый. 2018. №7(193). С. 8-19.

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

14. Grabner H., Grabner M., Bischof H. Real-time Tracking via On-line Boosting // Proceedings British Machine Vision Conference (BMVC). 2006. С. 47-56.

15. Babenko V. Visual Tracking with Online Multiple Instance Learning // Computer Vision and Pattern Recognition, CVPR 2009. IEEE Conference. 2009. С. 101-108.

16. Henriques J.F. High-Speed Tracking with Kernelized Correlation Filters // IEEE transactions on pattern analysis and machine intelligence. 2015. T.37(3). С. 301-314.

17. Ляшов М.В. Нейросетевая система отслеживания и распознавания объектов в видеопотоке // Современные наукоемкие технологии. 2018. № 12-1. С. 102-107.

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

19. Гуменюк М.М., Бровко А.В. Алгоритм отслеживания перемещений человека в видеопотоке с камер видеонаблюдения // Вестник Воронежского государственного университета. Серия: Системный анализ и информационные технологии. 2023. №3. С. 1-14.

20. danielgatis/rembg. URL: github.com/danielgatis/rembg (дата обращения: 11.10.2023).

21. Qin X., Zhang Z., Huang Ch., Dehghan M., Zaiane O.R., Jagersand M. U 2 -Net: Going Deeper with Nested U-Structure for Salient Object Detection // Pattern Recognition. 2020. V.106(11). URL: doi.org/10.1016/j.patcog.2020.107404 (дата обращения: 11.10.2023).

22. GitHub: [levindabhi/cloth-segmentation](https://github.com/levindabhi/cloth-segmentation). URL: github.com/levindabhi/cloth-segmentation (дата обращения: 11.10.2023).

References

1. Arkhipova E.A., Polyanichko K.S. Tomsk State University of Control Systems and Radio Electronics. 2020. № 7-2. pp. 69-72.

2. Nikitin D.V., Taranenko I.S., Kataev A.V. Inzhenernyj vestnik Dona (Rus), 2023, №8. URL: ivdon.ru/ru/magazine/archive/n7y2023/8531.

3. Mogilin K.A., Selishchev V.A. Proceedings of Tula State University. Technical sciences. 2020. №3. pp. 89-93.
 4. Moseva M.S., Kharrasov K.R., Nikitin D.V., Taranenko I.S., Kataev A.V. Inzhenernyj vestnik Dona, 2023, №8. URL: ivdon.ru/ru/magazine/archive/n8y2023/8643.
 5. Gumenyuk M.M. Problems of management in socio-economic and technical systems: collection of scientific articles (XVIII International Scientific and Practical Conference. Saratov, April 14-15, 2022) Saratov, 2022. pp. 126-132.
 6. Anstedt T., Keller I., Lutz H. Video Analytics: Myths and Reality 2nd ed., reprint. and add. Moscow: Security Focus. 2022. 186 p.
 7. Lukyanitsa A.A., Shishkin A.G. Digital video image processing. Moscow: I-Es-Es Press. 2009. 518 p.
 8. Minaev E.Yu., Kutikova V.V., Nikonorov A.V. Tracking objects in a video stream based on convolutional neural networks and fractal analysis: Proceedings (IV International Conference and youth school "Information technologies and nanotechnology" (ITNT-2018)) Samara: New Technique. 2018. pp. 2792-2798.
 9. Viola P., Jones M. Accepted Conference on Computer Vision and Pattern Recognition (CVPR 2001). 2001. V.1. pp. 511-518.
 10. Shelabin D.A., Sergeev S.L. Actual problems of humanities and natural sciences. 2013. №7-1. pp. 71-77.
 11. Artemov A.A., Kavalero M.V., Kuznetsov G.S. Bulletin of Perm State Technical University. Electrical engineering, information technology, control systems. 2011. №5. pp. 70-79.
 12. Zhivrin Ya.E., Alkzir N.B. Young scientist. 2018. №7(193). pp. 8-19.
 13. Pchelkin A.V., Nikiforov M.B. Proceedings of Tula State University. Technical sciences. 2020. №12. pp. 29-35.
-

14. Grabner H., Grabner M., Bischof H. Real-time Tracking via On-line Boosting: Proceedings (British Machine Vision Conference (BMVC)) 2006. pp. 47-56.
 15. Babenko B. Visual Tracking with Online Multiple Instance Learning: Proceedings (Computer Vision and Pattern Recognition, CVPR 2009. IEEE Conference) 2009. pp. 101-108.
 16. Henriques J.F. IEEE transactions on pattern analysis and machine intelligence. 2015. V.37(3). pp. 301-314.
 17. Lyashov M.V. Modern high-tech technologies. 2018. № 12-1. pp. 102-107.
 18. Pastushkov A.V. Method and algorithms for searching for an object in a video stream: abstract of the dissertation of the Candidate of Technical Sciences. Tomsk: TUSUR RAS, 2018. 135 pp.
 19. Gumenyuk M.M., Brovko A.V. Bulletin of the Voronezh State University. Series: System Analysis and Information Technology. 2023. №3. pp. 1-14.
 20. GitHub: [danielgatis/rembg](https://github.com/danielgatis/rembg). URL: github.com/danielgatis/rembg (accessed: 11.10.2023).
 21. Qin X., Zhang Z., Huang Ch., Dehghan M., Zaiane O.R., Jagersand M. Pattern Recognition. 2020. V.106 (11). URL: doi.org/10.1016/j.patcog.2020.107404 (accessed: 11.10.2023).
 22. GitHub: [levindabhi/cloth-segmentation](https://github.com/levindabhi/cloth-segmentation). URL: github.com/levindabhi/cloth-segmentation (accessed: 11.10.2023).
-