

# Телевизионные системы, передача и обработка изображений

УДК 004.93`12

## МЕТОДЫ ЛОКАЛИЗАЦИИ ОБЪЕКТОВ В ВИДЕОПОТОКЕ

### **Мареев Андрей Васильевич**

инженер кафедры физики и прикладной математики Муромского института (филиала) ФГБОУ ВО «Владимирский государственный университет имени А.Г. и Н.Г. Столетовых».

*E-mail:* [a.v.mareev@yandex.ru](mailto:a.v.mareev@yandex.ru)

### **Орлов Алексей Александрович**

доктор технических наук, доцент, заведующий кафедрой физики и прикладной математики Муромского института (филиала) ФГБОУ ВО «Владимирский государственный университет имени А.Г. и Н.Г. Столетовых».

*E-mail:* [alexeyalexorlov@gmail.com](mailto:alexeyalexorlov@gmail.com)

### **Рыжкова Мария Николаевна**

кандидат технических наук, доцент, доцент кафедры физики и прикладной математики Муромского института (филиала) ФГБОУ ВО «Владимирский государственный университет имени А.Г. и Н.Г. Столетовых».

*E-mail:* [masmash@mail.ru](mailto:masmash@mail.ru)

*Адрес:* 602264, Российская Федерация, Владимирская обл., г. Муром, ул. Орловская, д. 23.

*Аннотация:* Статья содержит обзор и анализ научных трудов в области локализации, обнаружения и отслеживания объектов в видеопотоке. Рассмотрена проблематика использования видеопотока для решения задач, недоступных при анализе одиночных кадров. Выполнен анализ методов обработки последовательности кадров цифровых изображений, которые позволяют локализовать объекты интереса на основе использования дополнительной информации на соседних кадрах. Исследованы точки зрения различных подходов к проблеме обработки видеопотока для обобщения и выделения конкретных групп методов. Проанализированные методы обработки видеопотока условно разделены на три группы. Рассмотрены широкие возможности комбинирования этих методов в различных областях деятельности. Определены перспективы развития методов данной области с точки зрения эффективности и быстроты действия.

*Ключевые слова:* локализация, изображения, видеопоток, комбинирование, методы обработки.

### **Введение**

В классических системах распознавания и автоматического ввода предполагается использование оцифрованного представления объекта, полученного сканированием или фотографией. При этом часто возникает необходимость решения задач локализации объектов (различимых яркостных областей) на изображении. Под локализацией понимается процесс определения областей объектов на изображениях по заданным параметрам [1]. Для локализации объектов на цифровых изображениях

существуют различные методы и алгоритмы, например методы, основанные на шаблонах (распознавание объектов на изображении на основе использования коэффициента корреляции или цветовой сегментации), методы с использованием контурных моделей (пространственное дифференцирование, преобразование Хафа), нейросетевые методы, метод Виолы — Джонса, метод опорных векторов, метод на основе получения нечёткого представления изображения [2], метод, основанный на детекторе Далала — Триггса [3], и другие [4].

При использовании видеокамеры для оцифровки образов распознаваемых объектов возникает возможность рассматривать видеопоток в целом, помимо отдельных фотографий или кадров. Видеопоток — это временная последовательность кадров (изображений) определённого формата [5].

Понятие локализации в видеопотоке тесно связано с понятиями обнаружения, распознавания и отслеживания объектов. Обнаружение (detection) — это установление наличия на изображении областей объектов с определёнными свойствами, а также, если объекты обнаружены, определение их координат на плоскости изображения [6]. Распознавание (recognition) — уточнение типов найденных объектов [7]. Отслеживание (tracking) — определение расположения объектов на каждом кадре в видеопотоке [7].

При наличии видеопотока в качестве цифрового образа объекта появляется возможность использовать гораздо больше визуальной информации [8]. Использование видеопотока позволяет решать задачи, недоступные для решения при анализе одиночной фотографии, когда внешние условия съёмки могут привести к тому, что распознаваемый объект перекрыт, недостаточно освещён, вышел из области съёмки, имеет неправильный ракурс съёмки или сильно искажён [9]. Примером является блик от протяжённого источника света, проявляющийся на глянцевой поверхности плоского объекта. В видеопотоке геометрическое положение снимаемого объекта, как правило, меняется между кадрами, блик также «сдвигается», что позволяет получить информацию о скрываемом объекте на другом кадре видеопотока. Существуют также важный класс объектов, локализация которых невозможна на одиночных снимках — к примеру, голографические элементы защиты, которые на единичных изображениях могут быть неотличимы от бликов или рисунков [10]. Рассматривая в качестве цифрового образа объекта не одиночное изображение, а видеопоток, появляется возможность уменьшить влияние ошибок за счёт

вариативности шума применительно к отдельным кадрам видеопотока, которой не обладают классические системы локализации и распознавания объектов. В таких условиях возникает задача выбора оптимальной стратегии комбинирования результатов покадрового анализа.

В разработке отечественных систем анализа видеопотока значительных успехов достигли школы таких научных деятелей как Петрова Е.П., Алпатова Б.А. и Арлазарова В.Л.

Научная школа Петрова Е.П. изучает выделение и сопровождение объектов на видеопотоке с неподвижным фоном с использованием цепей Маркова [10].

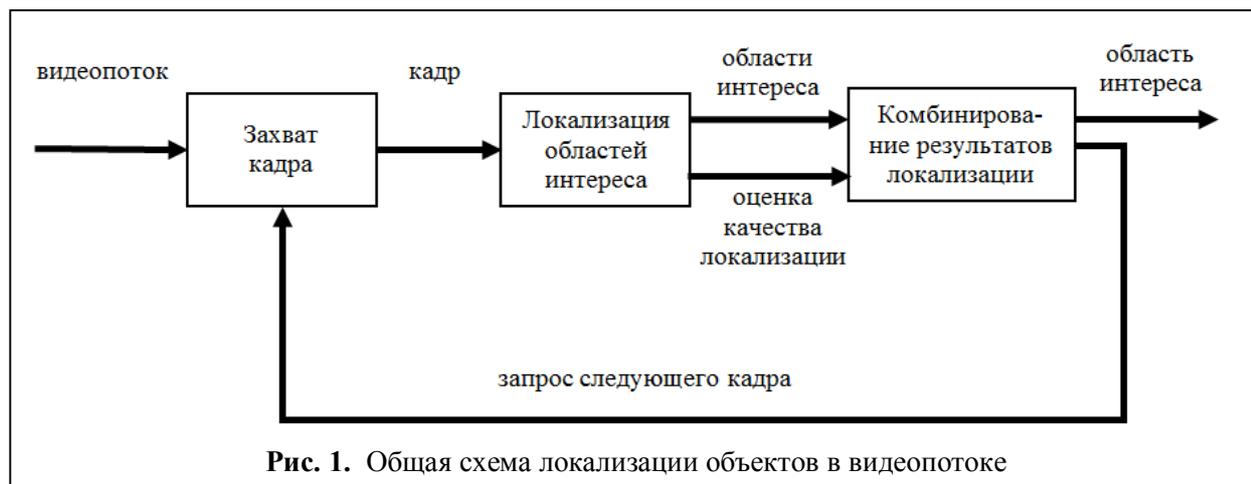
В работе [10] разработан метод выделения движущихся объектов в последовательности кадров на сравнительно неподвижном фоне на основе оценки их границ по вычисленной величине количества информации и произведён сравнительный анализ с другими алгоритмами, и представлены результаты, доказывающие высокую эффективность разработанного алгоритма.

Также значительных успехов в развитии отечественных систем анализа видеопотока добилась школа Алпатова. Алгоритмы школы Алпатова основаны на вычислении модели фона и области объекта интереса с целью локализации области объекта на неоднородном смещающемся фоновом изображении [11]. В статье [12] проведён обзор основных достижений научной школы.

Целью настоящей работы является обзор и анализ методов локализации, обнаружения и отслеживания объектов интереса в последовательности связанных цифровых изображений (в видеопотоке). Рассматриваются только методы, в которых выполняется комбинирование результатов локализации на нескольких предыдущих кадрах видеопотока (рис. 1).

Рассматриваемые методы и алгоритмы локализации, обнаружения и отслеживания в видеопотоке условно можно разделить на три группы:

- основанные на вычитании областей разных кадров;



- дообучающиеся на каждом кадре в случае удачной локализации;
- комбинирующие результаты анализа видеопотока различными методами и алгоритмами локализации объектов.

Методы находят применение для решения задач распознавания дорожных знаков, участников дорожного движения, номерных знаков, промышленной маркировки, текста, охранного видеонаблюдения и других задач.

### 1. Методы и алгоритмы, позволяющие выделить движущиеся объекты с помощью вычитания областей разных кадров

Существуют методы и алгоритмы, основанные на вычитании областей разных кадров, которые позволяют выделять движущиеся объекты в видеопотоке как на статичном, так и динамичном фоне.

Методы выделения движущихся объектов на статичном фоне описаны в работах [13–15]. Основным подходом, используемым в этих работах, является метод вычитания фона на основе Гауссовых распределений. Данный метод основан на адаптивном статистическом моделировании интенсивностей пикселей. Смесь из нескольких гауссиан используется для построения модели фона и детектирования движущихся объектов. При построении модели фона по данному методу считается, что для любого пикселя изображения известна история изменения его интенсивности на всех предыдущих кадрах. Методы, реализованные в рабо-

тах [13–15], заключаются в вычитании кадров изображения из эталонной модели изображения. Для этого из текущего кадра видео выполняется попиксельное вычитание интенсивностей предыдущих кадров видео и модели фонового изображения. Примерами практических задач, решаемых на основе данных методов локализации, является мониторинг потока машин [14], обнаружение движения людей в кадре [15], контроль за воспалительными процессами при совместном использовании тепловизора в операциях на сердце [13].

Обнаружение движущихся объектов с последующим отделением от динамического фона рассмотрено в статье [16]. В работе решается задача улучшения изображения (сглаживание перепадов яркости от кадра к кадру) на основе отслеживания движения объектов и фона, используя объединённый алгоритм глобального и локального оценивания оптического потока [17]. Этот алгоритм основан на трёх классических алгоритмах: Хорна — Шунка, Лукаса — Канаде и Брокса [18, 19], что обеспечивает возможность отслеживания как медленных, так и быстрых перемещений нескольких объектов.

### 2. Методы и алгоритмы, дообучающиеся на каждом кадре в случае удачной локализации

Методы и алгоритмы, дообучающиеся на каждом кадре в случае удачной локализации, можно разделить на три подгруппы:

- отслеживания предварительно выбранных объектов;
- обнаружения и отслеживания на основе оценки поведения объектов;
- локализации объекта в видеопотоке с использованием усовершенствованных алгоритмов обнаружения и отслеживания.

В первой подгруппе перед отслеживанием перемещения выделенного объекта, описанного в статьях [20, 21], обнаружение объекта интереса осуществляется вручную. Для отслеживания объекта в [26] рассматривается совместное использование алгоритмов TLD (Tracking–Learning–Detection) [22], CMT (Consensus–based Matching and Tracking) [23], Struck [24]. Авторы выполненной работы заключают, что совместное использование позволяет незначительно повысить точность отслеживания при равных значениях весовых коэффициентов вкладов выходных результатов алгоритмов при формировании итогового результата. В работе [21] применяется модифицированный алгоритм TLD, в котором для повышения точности был изменён компонент кратковременного отслеживания на фильтр частиц. Модификация алгоритма позволила повысить точность до алгоритма Struck и ухудшило быстродействие, но в условиях поворотов, изменения фоново-целевой обстановки, размеров объекта и окклюзий алгоритм, по заявлению авторов показывает хорошие результаты. Приведённые алгоритмы предлагается использовать для решения задач сопровождения объекта в видеопотоке, получаемом с камеры беспилотного летательного аппарата [28], сопровождения самолётов в полете, движущегося транспорта [21] и т. д.

Во второй подгруппе алгоритмы обнаружения и отслеживания на основе оценки поведения объектов, рассмотренные в работах [25, 26], используются в алгоритмах измерения расстояния до объекта [25] или оценки положения лица в пространстве [27]. Так, в работе [25] предлагается алгоритм для решения задачи оценки положения наблюдаемых объектов, основанный на алгоритме HCD (Harris Corner

Detector) [28] и оценки положения камеры на основе статичных объектов сцены. В методе отслеживания объектов на видео [27] применяется извлечение контурного представления объекта и обобщённое преобразование Хафа. Авторы в работе [26] предлагают использовать метод трекинга и описания объектов, основанный на использовании представления изображений и объектов как множества отдельных структурных элементов. Данный подход позволяет детектировать, отслеживать и стабильно удерживать обнаруженные объекты в разнообразных условиях.

Усовершенствованные алгоритмы третьей подгруппы, описанные в работах [29, 30], позволяют локализовать объекты в сложных условиях, в которых алгоритмы первой и второй подгруппы не справляются с поставленными задачами. Для этого применяется объединение нескольких зарекомендовавших себя методов. Например, объединение свёрточной нейронной сети и алгоритма отслеживания TLD в работе [29] позволяет корректировать алгоритм слежения, даже если объект был потерян.

В методе, представленном работой [31], обученная нейронная сеть YOLO объединена с методом OpenFace, который представляет исходное изображение в виде гистограммы направленных градиентов HOG (Histogram of Oriented Gradient) [30]. Метод применяется для отслеживания перемещения людей в контролируемом пространстве и отдельной их идентификации.

В статье [32] предложен статистический метод, позволяющий в реальном времени выделять объект на видеоизображениях и осуществлять его пространственную локализацию в последовательности видеок кадров. Для выделения объекта используются результаты его распознавания методом вероятностной фильтрации. Аппроксимация наиболее вероятного местоположения объекта эллипсом обеспечивает статистическую оценку координат центра объекта и позволяет описать траекторию движения объекта в относительных координатах.

В работе [33] решается задача поиска человека в видеопотоке за два этапа. На первом этапе на изображении для поиска выделяются всевозможные особенности изображения (углы, цветовые (яркостные) пятна, линии), каждая из которых затем сравнивается со всеми особенностями искомого объекта, используя коэффициент Бхаттачария, а затем отбрасываются особенности на изображении, которые не имеют ни одной схожести выше определённого порога. На втором этапе выполняется поиск области объекта среди всех отобранных подобластей изображения, которая имеет максимальное значение некоторого критерия схожести. В работе используется информация из цветовой гистограммы для локализации объекта при помощи метода Mean Shift [34].

Авторы в работе [35] предлагают выделять области объекта с использованием гистограмм для обнаружения объектов по шаблонам частей автомобиля. Распознавание шаблона реализуется при помощи разработанных специфических особенностей появления элементов транспортных средств, используемых человеком для определения частей автомобиля. Затем найденные таким образом классификаторы машинного обучения используются с Real Adaboost [30].

Примерами решаемых задач при помощи усовершенствованных алгоритмов являются идентификация личности, слежение за действиями определённых людей [31], отслеживание автомобилей в различных условиях с распознаванием номерных знаков [35] и т.д.

### **3. Методы и алгоритмы, комбинирующие результаты анализа видеопотока**

В статье [36] рассматривается задача поиска дорожных знаков с помощью камеры, установленной в транспортном средстве. Эта задача решается на основе использования модифицированного обобщённого преобразования Хафа и текущей скорости транспортного средства, которое прогнозирует перемещение области изображения с дорожным знаком на новых кадрах видеопотока. Данный метод улучшил

производительность системы, так как не требуется обрабатывать всю площадь новых кадров видеопотока. Таким образом, рассматриваемый алгоритм позволяет в реальном времени анализировать кадры высокого разрешения, что улучшает качество распознавания объектов интереса.

В работе [37] предлагается алгоритм отслеживания и обнаружения лица человека, позволяющий предоставить для распознавания качественные изображения. Достигается это путем анализа поступающего видеопотока с отбором наиболее информативных кадров методом корреляции. Также предлагается метод нормализации ракурса лица, основанный на детектировании особых точек и оценки ракурса POSIT (Pose from Orthography and Scaling with Iterations) [40].

Модель системы [39] предполагает пропуск кадров, полученных во время локализации объектов на ранее взятом изображении. Для локализации модель использует информацию об идеальной принадлежности каждого одиночного объекта к соответствующему классу. Выполняется поиск классифицирующей функции, которая максимизирует точность распознавания и интегрирует результаты, преобразующиеся в единый результат распознавания видеопоследовательности. Для решения задачи останова распознавания применяется экономическая модель, основанная на стоимости ввода корректного результата распознавания объекта, стоимости ввода ошибочного результата и стоимости распознавания одного изображения объекта.

Работа [40] посвящена комбинированию результатов покадрового распознавания символов как задаче коллективного принятия решения с множеством экспертов, уровни компетентности которых оцениваются при помощи предиктора достоверности результата распознавания. Наилучшие результаты комбинирования достигаются при помощи применения правил произведения и обобщённого голосования для подмножества нескольких самых компетентных экспертов.

В рамках работы [41] разработан метод поэтапного обнаружения маркировки груза для создания системы автоматической идентификации маркировки (САИМ). Система решает задачу наведения считывающей видеокамеры на маркировку груза исходя из положения груза в пространстве. На первом этапе в видеопотоке определяется текущее местоположение груза, на втором — маркировка, исходя из координат груза.

Алгоритм в работе [42] позволяет отслеживать местоположение предварительно выбранного объекта в видеопотоке и основан на объединении результатов работы алгоритмов отслеживания (трекер) и обнаружения (детектор) при помощи алгоритма корреляции шаблонов NCC (Normalized-Cross Correlation). Результатом работы трекера и детектора являются ограничительные окна. На основе NCC вычисляется параметр, который определяет величину подобия каждого полученного ограничительного окна и положительных примеров обучающей выборки. Ограничительное окно с наибольшим значением параметра подобия является результатом локализации объекта на текущем кадре и используется для обучения детектора с целью повешения точности обнаружения на следующих кадрах. Если ни трекер, ни детектор не выдают ограничительных окон, то считается, что объект потерян.

В работе [43] алгоритм локализации объектов состоит из нескольких этапов, в которых используется метод Виолы — Джонса и последовательное обучение. На первом этапе для оптимизации локализации объекта интереса выделяется область движения. На втором этапе при помощи сканирующего окна детектируются объекты интереса и определяются условия съёмки. Третий этап — это этап классификации, зависящий от условий съёмки, основан на совместном использовании трёх классификаторов: первый из них построен с помощью метода Виолы — Джонса и предварительно обучен, второй классификатор использует метод графов, третий классификатор построен на онлайн-обучении [44]. Классификатор онлайн-

обучения дообучается на обнаруженных объектах, полученных при плохих условиях съёмки.

Особенности и возможности рассмотренных методов и алгоритмов представлены в таблице 1 и таблице 2.

Наибольший интерес вызывают работы, в которых алгоритмы, обрабатывающие видеопоток с динамичным фоном, имеют возможность не только распознавать и идентифицировать объект интереса, но и позволяют локализовать его в случае потери при отслеживании. Подходы, предложенные в работах [37, 42], позволяют более достоверно распознавать объекты в видеопотоке.

В статьях [35, 36, 37, 42, 43] выполнена оценка разработанных методов и алгоритмов по показателю точности (таблица 3). Можно сделать вывод, что увеличение точности и уменьшение ошибок достигается за счёт иерархичного обнаружения. Т. е. сначала выделяют объекты, включающие область интереса, а затем осуществляется локализация подобъектов, одним из которых является область интереса.

### **Заключение**

В статье исследованы научные труды в области обнаружения и отслеживания объектов в видеопотоке. Приведённые методы и алгоритмы условно разделены на три группы: позволяющие выделить движущиеся объекты с помощью вычитания областей на разных кадрах, дообучающиеся на каждом кадре в случае удачной локализации и комбинирующие результаты локализации и отслеживания на каждом кадре. Основные особенности и возможности проанализированы и сведены в сравнительные таблицы.

Методы и алгоритмы, основанные на вычитании областей разных кадров, позволяют выделять в видеопотоке движущиеся объекты относительно стабильного фона. Но такие методы не подходят для локализации объектов, перемещающихся вместе с фоном.

**Таблица 1.** Особенности методов и алгоритмов

Ссылка на публикацию с описанием метода	Особенности методов и алгоритмов					Группа методов
	Динамично меняющийся фон	Вычитание фона	Использование более одного предыдущего кадра	Ручная предварительная локализация	Использование предобученной нейросети	
[13]	-	+	+	-	-	Основанные на вычитании областей разных кадров
[14]	-	+	+	-	-	
[15]	-	+	+	-	-	
[16]	+	+	+	-	-	
[20]	+	-	+	+	-	Дообучающиеся на каждом кадре в случае удачной локализации
[21]	+	-	+	+	-	
[25]	-	-	-	+	-	
[27]	-	-	-	+	+	
[26]	-	-	+	+	-	
[29]	-	-	-	+	+	
[31]	-	-	-	+	+	
[32]	+	-	-	+	-	
[33]	-	-	-	+	-	
[35]	+	-	-	-	+	
[36]	+	-	+	-	+	
[37]	+	-	+	+	-	Комбинирующие результаты анализа видеопотока различными методами и алгоритмами локализации объектов
[39]	-	-	+	-	-	
[40]	+	-	+	-	+	
[41]	+	-	-	-	-	
[42]	-	-	-	-	+	
[43]	-	-	-	+	-	

Методы и алгоритмы, дообучающиеся на каждом кадре, в случае удачной локализации основаны на совместном использовании алгоритмов TLD, CMT, Struck и свёрточных нейронных сетей. Данные методы позволяют решать задачи идентификации и отслеживания

сложных объектов — это задачи идентификации личности, слежения за действиями определённых людей, отслеживания автомобилей с распознаванием номерных знаков. Однако часто появляется проблема потери прослеживаемых

Таблица 2. Возможности методов и алгоритмов

Ссылка на публикацию с описанием метода	Возможности методов и алгоритмов						Группа методов
	Обучение во время работы	Обнаружение	Распознавание	Отслеживание	Идентификация	Локализация объекта в случае потери при отслеживании	
[13]	-	-	-	-	-	-	Основанные на вычитании областей разных кадров
[14]	-	+	-	-	-	-	
[15]	-	+	-	-	-	-	
[16]	-	+	-	-	-	-	
[20]	+	+	-	+	-	+	Дообучающиеся на каждом кадре в случае удачной локализации
[21]	+	+	-	+	-	+	
[25]	-	+	-	+	-	+	
[27]	-	+	-	+	-	-	
[26]	+	+	-	+	-	-	
[29]	+	+	+	+	-	+	
[31]	-	+	+	+	+	+	
[32]	-	+	+	+	-	+	
[33]	-	+	+	+	-	+	
[35]	-	+	+	+	+	-	
[36]	-	+	+	+	+	-	Комбинирующие результаты анализа видеопотока различными методами и алгоритмами локализации объектов
[37]	-	+	+	+	+	+	
[39]	-	+	+	-	-	+	
[40]	-	+	+	-	+	-	
[41]	-	+	+	-	+	+	
[42]	-	+	+	+	+	+	
[43]	+	+	+	-	+	-	

мых объектов интереса в случае недостаточной информативности их признаков.

В результате проведённого сравнения рассмотренных методов и алгоритмов можно сделать вывод, что комбинирование результатов анализа видеопотока является наиболее пер-

спективным для построения эффективных систем технического зрения. Такой подход прогнозирует поведение объекта, подстраивается под изменения видеопотока и проводит качественный анализ наиболее успешных кадров. Методы, использующие комбинирование ре-

**Таблица 3.** Заявленная точность методов и алгоритмов

Ссылка на публикацию с описанием метода	Точность, %
[35]	91,2
[36]	97,3
[37]	97,4
[42]	84
[43]	90

результатов анализа нескольких кадров видеопотока позволяют значительно увеличить точность локализации и более достоверно распознать объекты интереса в видеопотоке.

### Литература

1. Рубан И.В., Шитова О.В. Классификация методов обработки статических изображений для локализации объектов (областей «интереса») на них в системах технического зрения // Системы управления, навигации та зв'язку, випуск 3. 2009. № 11. С. 139–143.
2. Пугин Е.В., Жизняков А.Л. Алгоритмы обработки изображений для обнаружения объектов // Радиотехнические и телекоммуникационные системы. 2020. №2. С. 59–65.
3. Приоров А.Л., Хрящев В.В., Никитин А.Е., Матвеев Д.В. Анализ и модификация алгоритмов детектирования пешеходов на цифровых изображениях // Радиотехнические и телекоммуникационные системы. 2014. №4. С. 40–47.
4. Кудряшов П.П., Фоменков С.А. Гибридный алгоритм обнаружения человеческих лиц // Информационные технологии. 2007. № 10. С. 20–23
5. Забашта А.Ю., Мильченко М.А., Гришина Е.А. Обнаружение объектов в видеопотоке и построение их траектории, алгоритм центроидного отслеживания // Ростовский научный журнал. 2019. № 3. С. 238–248.
6. Сергеев В.В., Гашиников М.В., Мясников В.В. Обнаружение объектов на изображении [Электронный ресурс]. URL: <http://repo.ssau.ru/bitstream/Metodicheskie-ukazaniya/Obnaruzhenie-obektov-na-izobrazhenii-Elektronnyi-resurs-elektron-metod-ukazaniya-k-lab-rabote-4-53647/> (дата доступа 11.06.2021)
7. Борисов Е.С. О задаче поиска объекта на изображении. Часть 2: Применение методов машинного обучения [Электронный ресурс]. URL <http://mechanoid.kiev.ua/cv-image-detector2.html> (дата доступа: 11.06.2021)

8. Арлазаров В.В., Булатов К.Б., Усков А.В. Модель системы распознавания объектов в видеопотоке мобильного устройства // Труды ИСА РАН. Спецвыпуск. 2018. С. 73–82.

9. Арлазаров В.В., Жуковский А.Е., Кривцов В.Е., Николаев Д.П., Полевой Д.В. Анализ особенностей использования стационарных и мобильных малоразмерных цифровых видео камер для распознавания документов // Информационные технологии и вычислительные системы. 2014. № 3. С. 71–78.

10. Медведева Е.В., Карлушин К.А., Курбатова Е.Е. Метод выделения движущихся объектов в видеопотоке на основе оценки их границ // Машинное обучение и анализ данных. 2015. Т. 1, № 12. С. 1696–1705.

11. Алпатов Б.А., Бабаян П.В. Методы обработки и анализа изображений в бортовых системах обнаружения и сопровождения объектов // Цифровая обработка сигналов. 2006. № 2. С. 45–51.

12. Алпатов Б.А., Бабаян П.В. Технологии обработки и распознавания изображений в бортовых системах технического зрения // Вестник РГРТУ. 2017. № 60. С. 34–44.

13. Максименко В.Б., Шлыков В.В., Данилова В.А. Обнаружение тепловых неоднородностей для последовательности изображений в видео термограмм // Вісник НТУ «ХПІ». 2017. № 19(1241). С. 42–47.

14. Савостин А.А. Использование метода вычитания фона для обнаружения автотранспорта в видеопотоке // 11-я Международная молодежная научно-техническая конференция «Современные проблемы радиотехники и телекоммуникаций РТ-2015». 2015. С. 160.

15. 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.

16. Костенкова А.С., Грузман И.С. Выделение движущегося объекта в видеопотоке при помощи суперпиксельной сегментации и поля векторов движения // Современные проблемы радиоэлектроники. 2016. С. 20–23.

17. Liu C., Freeman W.T., Adelson E.H., Weiss Y. Human-assisted motion annotation // IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2008. Pp. 1–8.

18. Bruhn A., Weickert J., Schnörr C. Lucas/Kanade meets Horn/Schunck: combining local and global optic flow methods // International Journal of Computer Vision. 2005. № 61(3), Pp. 211–231.

19. Brox T., Bruhn A., Papenbergh N., Weickert J. High accuracy optical flow estimation based on a theory for warping // In European Conference on Computer Vision (ECCV). 2004. Pp. 25–36.

20. Павлов В.А., Завъялов С.В., Волвенко С.В. Анализ совместной работы методов сопровождения объектов в видеопотоке, получаемом с летательного аппарата // Анализ совместной работы. 2015. С. 75–79.
21. Занина М.А., Павлов В.А. Совместное применение алгоритма на основе фильтра частиц и tld для сопровождения объектов в видеопотоке // Неделя науки СПбПУ. 2018. С. 113–118.
22. Nebehay G., Pflugfelder R. Consensus-based Matching and Tracking of Keypoints for Object Tracking // Safety and Security Department AIT Austrian Institute of Technology 1220 Vienna. 2014.
23. Kalal Z., Mikolajczyk K., Matas J. Tracking-Learning-Detection // IEEE Transactions on pattern analysis and machine intelligence, 2010. Vol. 6, No. 1.
24. Hare S., Saffari A., Torr P.H.S. Struck: Structured Output Tracking with Kernels // Oxford Brookes University, Oxford, UK Sony Computer Entertainment Europe, London, UK 2012.
25. Кривоконь Д.С. Алгоритм оценки положений движущихся объектов с использованием рандомизации // Рандомизированные алгоритмы 2015. С. 3–19.
26. Азарков А.В. Выделение и отслеживание объектов на основе использования анализа движения // Штучний інтелект. 2015. С. 28–35.
27. Лошкарёв И.В., Демяненко Я.М. Применение обобщенного преобразования Хафа для поиска объектов в видеопотоке // Осенние математические чтения в Адыгее. 2015. С. 136–138.
28. Harris C., Stephens M. A combined corner and edge detector // In: Alvey Vision Conference. 1988. Vol. 15. Pp. 50.
29. Волков А.А., Войцеховский А.А., Зарницын А.Ю., Пякилла Б.И. Разработка алгоритма слежения за распознанными по видеопотоку объектами // XV Международная научно-практическая конференция студентов аспирантов и молодых ученых «Молодежь и современные информационные технологии». 2017. С. 132–133.
30. Dalal N., Triggs B. Histograms of Oriented Gradients for Human Detection. IEEE CVPR. 2005. Pp. 886–893.
31. Аксютин М.С., Гончарук С.Е. Распознавание объектов в видеопотоке при помощи алгоритма uolo и технологии openface // Научно-техническое творчество аспирантов и студентов. 2018. С. 208–211.
32. Попов А.В., Васильева И.К., Угарова А.С. Алгоритм сопровождения объектов в видеопотоке, регистрируемом с беспилотного летательного аппарата // Авиационно-космическая техника и технология. 2015. № 6 (123). С. 77–84.
33. Пастушков А.В., Калайда В.Т. Метод поиска человека на видеопотоке // Мир науки и инноваций. 2015. С. 39–42.
34. Comaniciu D., Meer P. Mean Shift: A Robust Approach Toward Feature Space Analysis // IEEE Transactions on Pattern Analysis and Machine Intelligence 24 (5). 2002. Pp. 603–619.
35. Мокишин В.В., Сайфудинов И.Р., Куртичников А.П., Шарнин Л.М. Распознавание образов транспортных средств на основе эвристических данных и машинного обучения // Вестник технологического университета. 2016. С. 130–137.
36. Якимов П.Ю. Распознавание дорожных знаков в реальном времени с использованием мобильного гпу // Информационные технологии и нанотехнологии. 2016. С. 977–982.
37. Небаба С.Г., Андреев С.Ю., Макаров М.А. Подготовка изображений лиц в видеопотоке к распознаванию и фильтрация неинформативных изображений // Средства и системы обработки и анализа данных. 2014. С. 1–8.
38. DeMenthon D.F. Model-based object pose in 25 lines of code // International Journal of Computer Vision. 1995. V. 15. I. 1–2. Pp. 123–141.
39. Arlazarov V.L., Loginov A.S., Slavin O.A. Characteristics of Optical Text Recognition Programs // Programming and Computer Software. 2002. Vol. 28, No. 3. Pp. 148–161.
40. Булатов К.Б. Выбор оптимальной стратегии комбинирования пок кадровых результатов распознавания символа в видеопотоке // Информационные технологии и вычислительные системы. 2017. № 3. С. 45–55.
41. Провоторов А.В., Орлов А.А. Методика поэтапного обнаружения маркировки слябов // Современные проблемы науки и образования. 2012. № 6. С. 1–5.
42. Ляшов М.В., Береза А.Н., Бабаев А.М., Коцюбинская С.А. Нейросетевая система отслеживания и распознавания объектов в видео потоке // Современные наукоемкие технологии. 2018. № 12. С. 102–107.
43. Бондарчук В.В., Шевченко Е.А. Исследование алгоритма детекции объектов интереса в видеопотоке // Проблемы искусственного интеллекта. 2015. № 1. С. 13–22.
44. Иванов Ю.С. Модифицированный алгоритм детекции лиц в видеопотоке и его программная реализация // Науковедение. 2014. № 3 (22). С. 1–26.

Поступила 11 июня 2021 г.

English

## METHODS OF OBJECT LOCALIZATION IN VIDEO STREAM

**Andrey Vasil'evich Mareev** — Engineer of Department of Physics and Applied Mathematics, Murom Institute (branch) "Vladimir State University named after A.G. and N.G. Stoletovs".

*E-mail:* [a.v.mareev@yandex.ru](mailto:a.v.mareev@yandex.ru)

**Alexey Alexandrovich Orlov** — Grand Dr. in Engineering, the Head of Department of Physics and Applied Mathematics, Murom Institute (branch) "Vladimir State University named after A.G. and N.G. Stoletovs".

*E-mail:* [alexeyalexorlov@gmail.com](mailto:alexeyalexorlov@gmail.com)

**Maria Nikolayevna Ryzhkova** — PhD, Associate Professor of Department of Physics and Applied Mathematics, Murom Institute (branch) "Vladimir State University named after A.G. and N.G. Stoletovs".

*E-mail:* [masmash@mail.ru](mailto:masmash@mail.ru)

*Address:* 602264, Russian Federation, Vladimir region, Murom, Orlovskaya St., 23.

*Abstract:* Object digitized representation is assumed to be used both in classical recognition systems and automatic input and which is obtained through scanning or photography. Upon that, there often arises necessity to solve object localization tasks (distinguishable brightness areas) in the image. However, there are tasks beyond of solving when analyzing a single photo when shooting external conditions may result in: identifiable object is blocked; it is not sufficiently lit; it is out of shooting coverage; it has the wrong shooting angle or is badly distorted. It is possible to consider the entire video stream besides individual photos or frames when using a camcorder to digitize images of identifiable objects. The purpose of this work is to review and analyze methods of localizing, detecting and tracking objects of interest in the sequence of linked digital images (in the video stream). Only those methods are considered where localization results are combined in several previous video stream frames. The article examines scientific works in the area of detecting and tracking objects in video stream. The above methods and algorithms are conditionally divided into three groups: enabling to single out moving objects by subtracting areas on different frames, which are learning on each frame in case of successful localization, and combining the results of localization and tracking on each frame. The main features and possibilities are analyzed and brought to comparative tables. Methods and algorithms based on subtracting areas of different frames enable to discern moving objects in video stream against stable background. However, such methods are not suitable for localizing objects moving along with the background. Methods and algorithms learning on each frame are based on co-utilization of TLD, CMT, Struck algorithms and convolution neural network in case of successful localization. These methods enable to solve tasks of identifying and tracking complex objects: these are tasks of person's identification, tracking actions of certain people, tracking cars and identifying their license plates. However, there often occurs the problem of losing traceable objects of interest in case of insufficient information content of their distinctive features. As a result of comparing the considered methods and algorithms the conclusion can be drawn that combination of analysis results of video stream is the most promising for making effective machine vision systems. This approach predicts object behavior, adjusts to video stream changes and performs qualitative analysis of the most successful frames. Methods that use combination of the analysis results of several frames of video stream enable to significantly enhance localization accuracy and more reliable recognition of objects of interest in video stream.

*Keywords:* localization, images, video stream, combination, processing methods.

### References

1. *Ruban I.V., Shitova O.V.* Classification of static image processing methods for localization of objects (areas of "interest") on them in technical vision systems. *Sistemi managlinnya, navigatsiya ta zvyazku*, Iss. 3. 2009. No. 11. Pp. 139–143.
2. *Pugin E.V., Zhiznyakov A.L.* Image processing algorithms for object detection. *Radiotekhnicheskie i telekommunikacionnye sistemy*. 2020. No. 2. Pp. 59–65.
3. *Priorov A.L., Khryashchev V.V., Nikitin A.E., Matveev D.V.* Analysis and modification of algorithms for detecting pedestrians on digital images. *Radiotekhnicheskie i telekommunikacionnye sistemy*. 2014. No. 4. Pp. 40–47.
4. *Kudryashov P.P., Fomenkov S.A.* Hybrid algorithm for detecting human faces. *Informacionnye tehnologii*. 2007. No. 10. Pp. 20–23
5. *Zabashta A.Yu., Milchenko M.A., Grishina E.A.* Detection of objects in the video stream and construction of their trajectory, centroid tracking algorithm. *Rostovskiy nauchnyi zhurnal*. 2019. No. 3. Pp. 238–248.

6. *Sergeev V. V., Gashnikov M. V., Myasnikov V. V.* Detection of objects in the image [Electronic source]. URL: <http://repo.ssau.ru/bitstream/Methodicheskie-ukazaniya/Obnaruzhenie-obektov-na-izobrazhenii-Elektronnyi-resurs-elektron-metod-ukazaniya-k-lab-rabote-4-53647/> (access date 11.06.2021)
7. *Borisov E. S.* About the task of searching for an object in the image. Part 2: Application of machine learning methods [Electronic source]. URL <http://mechanooid.kiev.ua/cv-image-detector2.html> (access date: 11.06.2021)
8. *Arlazarov V.V., Bulatov K.B., Uskov A.V.* Model of the object recognition system in the video stream of a mobile device. Proceedings of the ISA RAS. Special issue. 2018. Pp. 73–82.
9. *Arlazarov V.V., Zhukovsky A.E., Krivtsov V.E., Nikolaev D.P., Polevoy D.V.* Analysis of the features of the use of stationary and mobile small-sized digital video cameras for document recognition. Informacionnye sistemy i kompjuternye tehnologii. 2014. No. 3. Pp. 71–78.
10. *Medvedeva E.V., Karlushin K.A., Kurbatova E.E.* Method of selecting moving objects in a video stream based on the assessment of their boundaries. Mashinnoe obuchenie i analiz dannyh. 2015. Vol. 1, No. 12. Pp. 1696–1705.
11. *Alpatov B.A., Babayan P.V.* Methods of image processing and analysis in on-board systems for detecting and tracking objects. Digital signal processing. 2006. No. 2. Pp. 45–51.
12. *Alpatov B.A., Babayan P.V.* Technologies of image processing and recognition in on-board vision systems. Vestnik RGRTU. 2017. No. 60. Pp. 34–44.
13. *Maksimenko V.B., Shlykov V.V., Danilov V.A.* Detection of thermal irregularities for after-dovalidate the images in the video thermograms. Visnyk NTU "HP". 2017. № 19(1241). Pp. 42–47.
14. *Savostin A.A.* The method of background subtraction for the detection of vehicles in videopo-current. 11-th international youth scientific and technical conference "Modern problems of radio engineering and telecommunications RT-2015". 2015. P. 160.
15. *Zalivin A.N., Balabanova N.S.* Detection of moving objects by the method of background subtraction using a mixture of Gaussian distributions. Avtomatizirovannye tehnologii i proizvodstva. 2016. No. 3. Pp. 29–32.
16. *Kostenkova A.S., Gruzman I.S.* Selection of a moving object in a video stream using super-pixel segmentation and a field of motion vectors. Sovremennye problemy radioelektroniki. 2016. Pp. 20–23.
17. *Liu C., Freeman W.T., Adelson E.H., Weiss Y.* Human-assisted motion annotation // IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2008. Pp. 1–8.
18. *Bruhn A., Weickert J., Schnörr C.* Lucas/Kanade meets Horn/Schunck: combining local and global optic flow methods // International Journal of Computer Vision. 2005. № 61(3), Pp. 211–231.
19. *Brox T., Bruhn A., Papenberg N., Weickert J.* High accuracy optical flow estimation based on a theory for warping // In European Conference on Computer Vision (ECCV). 2004. Pp. 25–36.
20. *Pavlov V.A., Zavyalov S.V., Volvenko S.V.* Analysis of the joint work of object tracking methods in the video stream received from an aircraft. Analiz sovместnoj raboty. 2015. Pp. 75–79.
21. *Zanina M.A., Pavlov V.A.* Joint application of an algorithm based on a particle filter and a tld for tracking objects in a video stream. Nedelya nauki SPBPU. 2018. Pp. 113–118.
22. *Nebehay G., Pflugfelder R.* Consensus-based Matching and Tracking of Keypoints for Object Tracking // Safety and Security Department AIT Austrian Institute of Technology 1220 Vienna. 2014.
23. *Kalal Z., Mikolajczyk K., Matas J.* Tracking–Learning–Detection. IEEE Transactions on pattern analysis and machine intelligence, 2010. Vol. 6, No. 1
24. *Hare S., Saffari A., Torr P.H.S.* Structured Output Tracking with Kernels. Oxford Brookes University, Oxford, UK Sony Computer Entertainment Europe, London, UK 2012.
25. *Krivokon D.S.* Algorithm for estimating the positions of moving objects using randomization. Randomizirovannye algoritmy. 2015. Pp. 3–19.
26. *Agarkov A.V.* Selection and tracking of objects based on the use of motion analysis. Shtuchny intelekt. 2015. Pp. 28–35.
27. *Loshkarev I.V., Demyanenko Ya.M.* Application of the generalized Hough transform for searching for objects in a video stream. Osenniye matematicheskiye chteniya v Adygee. 2015. Pp. 136–138.
28. *Harris C., Stephens M.* A combined corner and edge detector. In: Alvey Vision Conference. 1988. Vol. 15. Pp. 50.
29. *Volkov A.A., Voitsekhovskiy A.A., Zarnitsyn A.Yu., Pyakillya B.I.* Development of an algorithm for tracking objects recognized by a video stream. XV International Scientific and Practical Conference of graduate students and young scientists "Youth and modern information technologies". 2017. Pp. 132–133.

30. *Dalal N., Triggs B.* Histograms of Oriented Gradients for Human Detection. IEEE CVPR. 2005. Pp. 886–893.
31. *Aksyutina M. S., Goncharuk S. E.* Object recognition in a video stream using the yolo algorithm and open-face technology. Scientific and technical creativity of graduate students and students. 2018. Pp. 208–211.
32. *Popov A.V., Vasilyeva I.K., Ugarov A.S.* Algorithm for tracking of objects in the video stream, registration with unmanned aircraft. Aviacionno-kosmicheskaya tekhnika i tekhnologiya. 2015. No. 6 (123). Pp. 77–84.
33. *Pastushkov A.V., Kalaida V.T.* The method of searching for a person on a video stream. Mir nauki i innovacij. 2015. Pp. 39–42.
34. *Moshkin V.V., Sayfutdinov I.R., Kirpichnikov A.P., Sharnin L.M.* Recognition of vehicle images based on heuristic data and machine learning. Vestnik tekhnologicheskogo universiteta. 2016. Pp. 130–137.
35. *Mokshin V.V., Sayfutdinov I.R., Kirpichnikov A.P., Sharnin L.M.* Recognition of vehicle images based on heuristic data and machine learning // Bulletin of the Technological University. 2016. pp. 130–137.
36. *Yakimov P.Yu.* Recognition of road signs in real time using a mobile GPU. Informacionnye tekhnologii i nanotekhnologii. 2016. Pp. 977–982.
37. *Nebaba S. G., Andreev S. Yu., Makarov M. A.* Preparation of images of faces in a video stream for recognition and filtering of uninformative images. Sredstva i sistemy obrabotki i analiza dannyh. 2014. Pp. 1–8.
38. *DeMenthon D.F.* Model-based object pose in 25 lines of code. International Journal of Computer Vision. 1995. Vol. 15. Iss. 1–2. Pp. 123–141.
39. *Arlazarov V.L., Loginov A.S., Slavin O.A.* Characteristics of Optical Text Recognition Programs. Programming and Computer Software. 2002. Vol. 28, № 3. Pp. 148–161.
40. *Bulatov K.B.* Choosing the optimal strategy for combining frame-by-frame results of character recognition in a video stream. Informacionnye tekhnologii i vychislitel'nye sistemy. 2017. No. 3. Pp. 45–55.
41. *Provotorov A.V., Orlov A.A.* Method of step-by-step detection of slab marking. Sovremennye problemy nauki i obrazovaniya. 2012. No. 6. Pp. 1–5.
42. *Lyashov M.V., Bereza A.N., Babaev A.M., Kotsyubinskaya S.A.* Neural network system for tracking and recognizing objects in a video stream. Sovremennye naukoemkie tekhnologii. 2018. No. 12. Pp. 102–107.
43. *Bondarchuk V.V., Shevchenko E.A.* Investigation of the algorithm for detecting objects of interest in a video stream. Problemy iskusstvennogo intellekta. 2015. No. 1. Pp. 13–22.
44. *Ivanov Yu.S.* A modified algorithm for detecting faces in a video stream and its software implementation. Naukovedenie. 2014. No. 3 (22). Pp. 1–26.