

УДК 004.93

ПРИКЛАДНЫЕ ЗАДАЧИ ИНТЕЛЛЕКТУАЛЬНОЙ ОБРАБОТКИ ВИДЕОПОТОКА

Карпик С.Э., магистрант гр.55624

*Белорусский государственный университет информатики и радиоэлектроники,
Институт информационных технологий,
г. Минск, Республика Беларусь*

Парамонов А.И. – канд. техн. наук, заведующий кафедрой ИСиТ

Аннотация. Рассматривается проблема интеллектуальной обработки видеопотока в условиях роста объемов видеоданных и увеличения нагрузки на операторов. Проанализированы ограничения традиционных систем видеонаблюдения и обоснована необходимость перехода к автоматизированному анализу визуальной информации видеопотока. Описана архитектура современных систем видеоаналитики, включая этапы предварительной обработки, детекции, классификации и трекинга объектов. Рассмотрены методы компьютерного зрения и машинного обучения, а также основные прикладные области применения интеллектуальных систем видеонаблюдения.

Ключевые слова. Интеллектуальная обработка видеопотока, видеоаналитика, компьютерное зрение, детекция объектов, трекинг объектов, нейронные сети, глубокое обучение, OpenCV.

Введение. В статье рассматривается проблема интеллектуальной обработки видеопотока и её применение для решения прикладных задач анализа визуальной информации. В современном мире системы видеонаблюдения и видеомониторинга широко используются в различных областях — от обеспечения общественной безопасности до промышленного контроля и управления транспортными потоками. Количество установленных камер видеонаблюдения ежегодно увеличивается, что приводит к резкому росту объемов видеоданных, требующих обработки и анализа.

Традиционные системы видеонаблюдения в основном выполняют функцию записи и хранения видеоматериалов. Однако такой подход имеет существенные ограничения.

Во-первых, анализ видеозаписей часто осуществляется человеком-оператором, что требует значительных трудовых ресурсов.

Во-вторых, человек способен эффективно контролировать лишь ограниченное количество видеопотоков. При длительной монотонной работе внимание оператора снижается, что повышает вероятность пропуска значимых событий.

В-третьих, увеличение числа камер делает ручной анализ видеоданных практически невозможным.

В последние годы наблюдается стремительный рост количества систем видеонаблюдения. По различным оценкам, в крупных городах количество установленных камер исчисляется десятками и сотнями тысяч. В результате ежедневно формируются терабайты видеоданных, которые требуют хранения, обработки и анализа. Обработка таких объемов информации традиционными методами становится практически невозможной. Это приводит к формированию нового класса систем — интеллектуальных систем видеоаналитики, способных автоматически анализировать визуальные данные и извлекать из них полезную информацию.

В условиях роста количества видеопотоков возникает необходимость автоматизации их анализа. Интеллектуальная обработка видеопотока представляет собой совокупность методов компьютерного зрения и машинного обучения, направленных на автоматическое обнаружение объектов, анализ поведения, распознавание событий и выявление аномалий в видеоданных [1]. Такие системы могут работать как в режиме реального времени, так и при обработке архивных записей.

Развитие технологий компьютерного зрения привело к появлению эффективных алгоритмов анализа изображений и видео. Особенно значительный прогресс был достигнут с внедрением методов глубокого обучения, в частности сверточных нейронных сетей. Эти методы позволяют автоматически извлекать признаки из изображений и обеспечивают высокую точность распознавания объектов даже в сложных условиях наблюдения.

Архитектурные решения современных систем мониторинга видеопотока.

Современные системы интеллектуального анализа видео строятся по многоуровневой архитектуре, включающей несколько этапов обработки данных. Такой подход позволяет эффективно обрабатывать большие объемы видеоданных и распределять вычислительную нагрузку между различными компонентами системы.

На первом этапе выполняется захват видеопотока. Источниками данных могут выступать IP-камеры, видеорегистраторы, потоковые протоколы передачи данных (например, RTSP), а также архивные видеозаписи. В некоторых случаях используются специализированные камеры с поддержкой встроенной аналитики.

После получения видеопотока выполняется декодирование кадров и их предварительная обработка. Данный этап играет важную роль, поскольку позволяет уменьшить влияние шумов, улучшить качество изображения и подготовить данные для последующего анализа. Предварительная обработка может включать фильтрацию шумов, коррекцию яркости и контрастности, стабилизацию изображения, нормализацию цветового пространства, масштабирование изображений.

Следующим этапом является выделение областей интереса на изображении. Для этого применяются методы обнаружения движения, вычитания фона и поиска контуров. Основная задача данного этапа заключается в том, чтобы сократить объём информации, передаваемой на последующие стадии обработки.

Полученные области интереса передаются на следующий уровень обработки — системы распознавания объектов. На данном этапе используются алгоритмы компьютерного зрения, позволяющие определить наличие объектов на изображении и определить их положение. После обнаружения объектов выполняется их классификация.

После обнаружения объектов на отдельных кадрах возникает задача их сопоставления между последовательными кадрами видеопотока. Этот процесс называется трекингом объектов. Основной целью трекинга является сохранение идентификатора объекта на протяжении всей последовательности кадров и определение его траектории движения.

Одним из распространённых подходов является использование фильтра Калмана, который позволяет прогнозировать положение объекта на следующем кадре на основе предыдущих наблюдений. Для сопоставления обнаруженных объектов между кадрами применяется алгоритм венгерского назначения.

В современных системах широко используются алгоритмы SORT (Simple Online Realtime Tracking) и Deep SORT. Алгоритм Deep SORT дополняет классический подход использованием нейросетевых признаков внешнего вида объекта, что позволяет уменьшить количество ошибок сопоставления.

На заключительном этапе выполняется анализ событий и интеграция результатов в прикладные системы. Полученная информация может использоваться для формирования сигналов тревоги, записи событий в базу данных, управления оборудованием или визуализации результатов для оператора.

Предварительная обработка является важным этапом интеллектуального анализа видеопотока, поскольку позволяет повысить точность последующих алгоритмов обработки. На практике используются различные методы обработки изображений, позволяющие выделять информативные признаки сцены.

Одним из наиболее распространённых методов является вычитание фона. Данный метод основан на построении модели статического фона сцены и сравнении текущего кадра с этой моделью. Если на изображении появляются новые объекты, они рассматриваются как движущиеся элементы.

В библиотеке OpenCV реализованы алгоритмы MOG и MOG2, основанные на гауссовых моделях смеси. Эти алгоритмы позволяют адаптивно обновлять модель фона и учитывать постепенные изменения освещения.

Другим важным методом является поиск контуров. Контуров позволяют определить границы объектов на изображении и выделить отдельные области интереса. После обнаружения контуров выполняется их фильтрация по различным параметрам, таким как площадь, форма или положение.

Также широко применяются морфологические операции обработки изображений. К ним относятся операции расширения, эрозии, открытия и закрытия. Эти методы позволяют устранить шумы, заполнить небольшие разрывы в объектах и улучшить качество сегментации изображения. Дополнительно могут использоваться методы компенсации дрожания камеры, а также алгоритмы подавления теней. Эти методы особенно важны для систем видеонаблюдения, установленных на открытых пространствах.

В последние годы основным направлением развития компьютерного зрения стало использование методов глубокого обучения. Сверточные нейронные сети продемонстрировали высокую эффективность в задачах анализа изображений и видео.

Одной из наиболее известных архитектур является YOLO (You Only Look Once) [2]. Данный алгоритм выполняет обнаружение объектов на изображении за один проход нейронной сети, что обеспечивает высокую скорость обработки. Благодаря этому YOLO широко применяется в системах анализа видео в реальном времени [3].

Другим популярным алгоритмом является SSD (Single Shot Detector). Он также относится к методам одношаговой детекции объектов и обеспечивает хорошее соотношение скорости и точности.

Для более сложных задач сегментации используется архитектура Mask R-CNN. Она позволяет не только обнаружить объект, но и определить его точную форму на уровне пикселей.

Помимо детекции объектов важной задачей является их отслеживание во времени. Для этого используются алгоритмы трекинга, такие как Kalman Filter, SORT, Deep SORT. Эти алгоритмы позволяют связывать объекты между кадрами видеопотока и анализировать их движение.

Прикладные задачи интеллектуальной обработки видеопотока

Современные системы интеллектуального анализа видеопотока находят применение в различных областях.

Одним из наиболее распространённых применений является анализ дорожного движения. Системы видеоаналитики позволяют обнаруживать транспортные средства, подсчитывать их количество, анализировать плотность движения и выявлять нарушения правил дорожного движения.

Такие системы могут фиксировать пересечение стоп-линии, движение по выделенной полосе или въезд в запрещённую зону.

В промышленности интеллектуальные системы анализа видео используются для контроля производственных процессов. Камеры устанавливаются над конвейерными линиями и автоматически отслеживают перемещение продукции.

Система может подсчитывать количество произведённых изделий, обнаруживать дефекты и фиксировать отклонения в технологическом процессе.

В системах безопасности алгоритмы компьютерного зрения используются для обнаружения подозрительных действий, анализа поведения людей и выявления потенциальных угроз. Например, система может обнаружить проникновение на охраняемую территорию, обнаружить оставленный предмет или зафиксировать агрессивное поведение.

В концепции «умных» зданий видеоаналитика используется для анализа загруженности помещений, управления освещением и оптимизации использования ресурсов. Система может определять количество людей в помещении, анализировать потоки посетителей и автоматически управлять инженерными системами здания.

Ещё одной важной задачей является поиск событий в архивных видеозаписях. Использование алгоритмов компьютерного зрения позволяет автоматически индексировать видеоданные и выполнять поиск по заданным параметрам. Это значительно ускоряет анализ видеоматериалов и упрощает работу с большими архивами.

Анализ современных систем интеллектуальной обработки видеопотока показывает, что наилучшие результаты достигаются при комбинировании различных методов обработки данных. Классические методы компьютерного зрения обладают высокой скоростью работы и низкими требованиями к вычислительным ресурсам. Они хорошо подходят для предварительной обработки данных и выделения областей интереса. Однако такие методы имеют ограниченные возможности и плохо справляются с задачами сложного распознавания объектов.

Методы глубокого обучения обеспечивают более высокую точность распознавания и способны работать в сложных условиях. Они устойчивы к изменениям освещения, частичным перекрытиям объектов и сложной геометрии сцены. Тем не менее использование нейросетевых моделей требует значительных вычислительных ресурсов и наличия обучающих данных. Поэтому современные системы видеоаналитики обычно используют гибридную архитектуру, в которой классические алгоритмы выполняют предварительную обработку данных, а нейронные сети выполняют сложный анализ изображений.

Интеллектуальный анализ видеопотока требует значительных вычислительных ресурсов, особенно при обработке видео высокого разрешения в режиме реального времени. Для ускорения вычислений широко используются графические процессоры (GPU), которые позволяют выполнять параллельную обработку данных.

Одним из распространённых решений являются специализированные вычислительные платформы для edge-обработки, такие как NVIDIA Jetson. Эти устройства позволяют выполнять нейросетевой анализ видео непосредственно на уровне камеры или локального вычислительного узла, что снижает задержки передачи данных и уменьшает нагрузку на центральный сервер.

Заключение.

В статье была рассмотрена проблема интеллектуальной обработки видеопотока и проанализированы основные методы и технологии, применяемые в системах видеоаналитики. Показано, что современные системы анализа видео постепенно переходят от пассивного хранения видеоданных к активному анализу происходящих событий. Использование методов компьютерного

зрения и машинного обучения позволяет значительно повысить эффективность обработки видеоданных и автоматизировать решение многих прикладных задач.

Одним из ключевых направлений развития является применение более современных архитектур глубокого обучения. В последние годы активно исследуются трансформерные модели (Vision Transformers и Video Transformers), которые позволяют учитывать пространственно-временной контекст видеоданных. В отличие от классических сверточных нейронных сетей, такие архитектуры способны анализировать взаимосвязи между объектами на разных участках изображения и учитывать динамику сцены во времени. Это особенно важно для задач анализа поведения людей, обнаружения сложных событий и прогнозирования развития ситуации.

Другим перспективным направлением является использование методов self-supervised learning и semi-supervised learning. Традиционные модели машинного обучения требуют больших объемов размеченных данных для обучения, получение которых является трудоёмким и дорогостоящим процессом. Использование методов самообучения позволяет обучать модели на больших объемах неразмеченных видеоданных, что существенно снижает зависимость от ручной разметки и ускоряет разработку новых систем видеоаналитики.

Значительное внимание уделяется развитию технологий edge computing. В классической архитектуре систем видеонаблюдения видеопоток передаётся на центральный сервер, где выполняется его обработка. Однако при большом количестве камер такая схема приводит к высокой нагрузке на сеть и вычислительные ресурсы серверов. Использование периферийных вычислительных устройств (edge-устройств) позволяет выполнять часть обработки непосредственно на уровне камеры или локального вычислительного узла. Такой подход снижает задержки обработки данных и повышает масштабируемость системы.

Ещё одним перспективным направлением является разработка более компактных и энергоэффективных нейросетевых моделей. В условиях использования интеллектуальной видеоаналитики на мобильных устройствах, беспилотных системах и встроенных платформах возникает необходимость оптимизации моделей по размеру и вычислительной сложности. Для этого применяются методы квантования, прунинга нейронных сетей и knowledge distillation, позволяющие существенно уменьшить размер моделей без значительного снижения точности.

Перспективным направлением является также развитие методов автоматического анализа поведения и обнаружения аномалий. Такие системы способны выявлять нестандартные или потенциально опасные ситуации без заранее заданных сценариев. Это особенно актуально для систем безопасности, мониторинга общественных пространств и промышленного контроля.

Развитие интеллектуальных систем анализа видеопотока позволит существенно повысить уровень автоматизации различных отраслей экономики и создать новые возможности для использования видеоданных в качестве источника аналитической информации.

Список использованных источников:

1. Szeliski, R. *Computer Vision: Algorithms and Applications* [Электронный ресурс] / R. Szeliski. – Режим доступа: <https://szeliski.org/Book/> – Дата доступа: 10.04.2026.
2. *Small Object Detection with YOLO: A Performance Analysis Across Model Versions and Hardware* [Электронный ресурс]. – Режим доступа: <https://arxiv.org/html/2504.09900v1> – Дата доступа: 10.04.2026.
3. *YOLO Evolution: A Comprehensive Benchmark and Architectural Review of YOLOv12, YOLO11, and Their Previous Versions* [Электронный ресурс]. – Режим доступа: <https://arxiv.org/html/2411.00201v2> – Дата доступа: 10.04.2026.

UDC 004.93

APPLIED TASKS OF INTELLIGENT VIDEO STREAM PROCESSING

Karpik S.E.

*Institute of Information Technologies of the Belarusian State University of Informatics and Radioelectronics,
Minsk, Republic of Belarus*

Paramonov A.I. – Candidate of Technical Sciences, Associate Professor

Annotation. The problem of intelligent video stream processing is considered under conditions of growing volumes of video data and increasing operator workload. The limitations of traditional video surveillance systems are analyzed, and the necessity of transitioning to automated analysis of visual information from video streams is substantiated. The architecture of modern video analytics systems is described, including stages of preprocessing, detection, classification, and object tracking. Methods of computer vision and machine learning are examined, as well as the main application areas of intelligent video surveillance systems.

Keywords. Neural Intelligent video stream processing, video analytics, computer vision, object detection, object tracking, neural networks, deep learning, OpenCV.