

Пеньковой М.А. (Университет ИТМО)
Научный руководитель – к.т.н., доцент Ключев А.О.
(Университет ИТМО)

В данной работе выполнен обзор методов компьютерного зрения, которые применяются для решения задачи определения траектории движения объектов на видео. Разработан прототип системы видеонаблюдения на основе нейронной сети YOLOv4 и алгоритма Deep SORT, проведено исследование работы прототипа на тестовом наборе данных.

Введение. В настоящее время активно разрабатываются интеллектуальные системы видеонаблюдения, способные выполнять анализ поступающей с камер информации без участия человека. Такие системы применяются для обеспечения безопасности, контроля закрытых территорий, поиска и отслеживания перемещений людей, анализа посещаемости торговых точек, используются на промышленных предприятиях и в умных домах. Для эффективного выполнения этих задач интеллектуальные системы видеонаблюдения должны уметь определять траекторию движения объектов, то есть не только распознавать их на текущем кадре, но и анализировать объекты с учетом информации, полученной с предыдущих кадров. Существующие решения данной проблемы основаны на использовании методов вычисления оптического потока, техники сдвига среднего значения, либо алгоритмов машинного обучения.

Основная часть. Оптический поток – это модель видимого движения объектов, вызванного относительным перемещением между сценой и наблюдателем. Для определения сдвига точек используется информация об изменении интенсивности пикселей двух последовательных кадров. Применяя метод Лукаса-Канаде, можно получить уравнение для скорости отслеживаемых пикселей, решение которого, позволяет определить траекторию движения объекта в видеопотоке. В основе методов вычисления оптического потока лежит ряд предположений о том, что интенсивность пикселей, меняющих свое расположение, остается постоянной, движения выполняются плавно, а соседние пиксели, принадлежащие одному объекту, перемещаются согласовано. Однако эти предположения накладывают ограничения на практическое применение данных методов.

Техника сдвига среднего значения основывается на анализе набора свойств объекта для определения местоположения максимума плотности вероятности. Для обнаружения объектов используются его признаки, обычно применяется цветовая гистограмма. На ее основе строится функция плотности вероятности, называемая картой достоверности, которая присваивает каждой точке кадра вероятность, того, что ее цвет встречался в объекте на предыдущем кадре. Рассчитав карту достоверности для текущего изображения на основе цветовой гистограммы объекта и применив технику сдвига среднего значения, можно найти максимум карты достоверности, и получив текущее положение объекта, отследить его перемещение.

Алгоритм SORT основан на применении фильтра Калмана и венгерского алгоритма. С помощью методов компьютерного зрения происходит распознавание объектов, определение их размеров, скоростей и положения, для последующего анализа. Для уменьшения влияния шумов применяется фильтр Калмана. Данный фильтр является рекурсивным алгоритмом, который использует серию зашумленных и неполных измерений переменных, наблюдаемых в определенный промежуток времени, и выполняет оценку совместного распределения вероятностей по переменным для каждого периода времени из промежутка, предсказывая значения неизвестных переменных, которые получаются точнее, значений, основанных на единственном измерении. Для повышения точности работы алгоритма SORT применяется расстояние Махаланобиса, которое в отличие от евклидова расстояния, учитывает дисперсию

и линейную зависимость между исследуемыми переменными. Чтобы связать объекты, распознанные на новом кадре с объектами, обнаруженными на предыдущих кадрах, используется венгерский алгоритм.

Алгоритм Deep SORT расширяет исходный алгоритм SORT, добавляя анализ внешнего вида отслеживаемых объектов. Использование признаков внешнего вида позволяет устранить визуальную потерю объектов, вызванную окклюзиями, изменениями точек обзора и шумами, повышая точность определения траектории их движения.

Для реализации прототипа системы видеонаблюдения было принято решение использовать нейронную сеть YOLOv4 для обнаружения объектов и алгоритм Deep SORT для отслеживания их движения. Обучение нейронной сети происходило на наборе данных COCO, содержащем объекты разных классов. Тестирование разработанного прототипа системы видеонаблюдения показало приемлемое качество и время распознавания, пригодное для отслеживания движения объектов в режиме реального времени.

Выводы. В результате данного исследования был выполнен обзор методов компьютерного зрения, применяемых для определения траектории движения объектов в видеопотоке. Был разработан прототип системы видеонаблюдения, с использованием алгоритма Deep SORT и нейронной сети YOLOv4. Разработанный прототип может применяться для распознавания объектов и слежения за их перемещением в видеопотоке.

Пеньковой М.А. (автор)

Подпись

Ключев А.О. (научный руководитель)

Подпись