

УДК 004.932.72

## ПОИСК И СОПРОВОЖДЕНИЕ ОДНОГО ИЛИ НЕСКОЛЬКИХ ОБЪЕКТОВ С ПОМОЩЬЮ АНАЛИЗА АЭРОФОТОСНИМКОВ

Безбородов А.К. (Санкт-Петербургский политехнический университет Петра Великого)

Научный руководитель – ведущий инженер Павлов В.А.

(Санкт-Петербургский политехнический университет Петра Великого)

Данный доклад посвящен описанию алгоритма обнаружения, распознавания и сопровождения движущихся объектов в условиях аэрофотосъемки с использованием подвижной камеры на базе мобильной вычислительной платформы. В отличие от большого числа существующих алгоритмов, ключевой особенностью разработанного подхода является возможность устанавливать взаимно-однозначное соответствие между одинаковыми объектами в кадре.

**Введение.** Важнейшими задачами в области компьютерного зрения являются обнаружение, распознавание и сопровождение одного или нескольких объектов на видеопоследовательности. Спектр применения таких алгоритмов чрезвычайно широк и включает в себя различные сферы человеческой деятельности: промышленность (контроль качества производимых объектов), видеоаналитика (подсчет посетителей), мониторинг местности (фиксация нарушений, безопасность на производстве). Для подобных задач требуется не только производить своевременную оценку траектории движения обнаруженных объектов, но также отслеживать и сопоставлять уникальные объекты от кадра к кадру.

Большинство имеющихся на данный момент методов, направленных на решение задач по обнаружению и отслеживанию движущихся объектов, демонстрируют свою эффективность в случае, если изображение было получено с фиксированно расположенной камеры. Существующие алгоритмы, применение которых возможно в случае съемки с подвижной камеры, демонстрируют свою эффективность при запуске на компьютерных устройствах, обладающих достаточно большой вычислительной мощностью. Таким образом, основное требование к разработанному подходу заключается в максимальной точности обнаружения и отслеживания объектов в условиях сложной фоновой-целевой обстановки.

**Основная часть.** В современных алгоритмах имеется четкое различие между первоначальным обнаружением, последующим отслеживанием и сопоставлением движущихся объектов. Прежде всего, необходимо определить эффективный метод их поиска на каждом из кадров входной видеопоследовательности. В большом количестве имеющихся подходов для нахождения объектов требуется первоначальное обнаружение переднего плана, что позволяет отделить движущиеся в кадре объекты от неподвижных, относящихся к заднему плану. Движущимися объектами считают те области переднего плана, для которых возможно произвести оценку их местоположения на следующем кадре.

Основными способами отделения движущихся объектов переднего плана от заднего являются: вычитание фона, оптический поток и различные вероятностные методы. На практике такие способы поиска объектов корректно работают лишь в случае неподвижно закрепленной камеры, поэтому их использование в условиях аэрофотосъемки оказывается невозможным.

Подвижная камера и динамичный фон создают большое количество трудностей, поэтому для нахождения объектов, их классификации и сопровождения был выбран метод с использованием сверточных нейронных сетей. В случае применения такого подхода выделять передний план в изображениях не требуется, так как с выхода нейронной сети формируется изображение с выделенными интересующими нас объектами и указанием их классов.

Для данной работы была выбрана предварительно обученная на наборе изображений, полученных с помощью аэрофотосъемки, сверточная нейронная сеть Tiny-YOLOv2. В ней, по сравнению с обычной версией, с целью существенного уменьшения вычислительных затрат и выигрыша в количестве кадров в секунду при обработке видеопоследовательности, было

уменьшено количество сверточных слоев без ощутимых потерь в точности обнаружения объектов. Запускать обученную нейронную сеть требуется для каждого из кадров входной видеопоследовательности, то есть в режиме реального времени.

Финальным этапом является сопоставление уникальных объектов. Соответствующая обработка выполняется с использованием алгоритма Куна-Манкреса и сравнением ключевых точек обнаруженных объектов.

При использовании алгоритма Куна-Манкреса требуется определить матрицу значений, которая, с помощью последовательных линейных алгебраических преобразований, приводится к соответствующему оптимальному виду. В качестве элементов матрицы были выбраны значения метрики евклидова расстояния между центрами всех ограничивающих прямоугольников текущего и предыдущего кадров. При минимальном евклидовом расстоянии между двумя объектами на последовательных кадрах можно говорить об идентификации одного и того же объекта.

При тестировании данного подхода были получены достаточно точные результаты сопоставления движущихся объектов на видеопоследовательности от кадра к кадру. Однако в случае повторного появления одно и того же объекта в кадре, алгоритм Куна-Манкреса устанавливает для обнаруженного объекта новый индекс, что предоставляет ложные сведения об отслеживаемом объекте.

Обработка таких проблемных ситуаций производится с помощью метода сравнения ключевых точек. Ключевой называют точку, которая является уникальной в определенной области вокруг себя и может быть сопоставлена с соответствующей ей точкой на другом изображении.

В каждом из ограничивающих прямоугольников, полученных с выхода нейронной сети и содержащих объект для дальнейшего отслеживания, применяется программный метод по нахождению ключевых точек с использованием функций библиотеки OpenCV. В данной библиотеке существует множество алгоритмов поиска ключевых точек, такие как FAST, STAR, SIFT, SURF, BRISK, ORB и многие другие. При исследовании поведения BRISK ключевых точек была выявлена их относительная устойчивость при резких смещениях и поворотах камеры, что и обусловило выбор такого метода обнаружения ключевых точек.

Принцип обработки выделенных признаков объектов заключается в формировании массива из 10 ключевых точек для каждого движущегося объекта, который постоянно обновляется от кадра к кадру. При выходе объекта из кадра, а затем повторном появлении последние обнаруженные ключевые точки всех объектов, отслеживание которых не производится в текущий момент времени, сравниваются с ключевыми точками нового обнаруженного объекта. В случае большого числа совпадений предполагается, что объект, ранее покинувший кадр, вернулся в поле зрения камеры.

**Выводы.** Разработанный алгоритм может применяться в задачах сопровождения определенных движущихся объектов в ходе их перемещения при съемке с борта беспилотного летательного аппарата, а также для мониторинга местности в режиме реального времени и контроля периметра охраняемых объектов.

Запуск всего алгоритма планируется осуществлять на базе мобильной вычислительной платформы NVIDIA Jetson Nano. Данный модуль позволяет обеспечить требуемую производительность и эффективность, необходимую для запуска различного рода приложений на базе искусственного интеллекта в условиях ограниченных габаритами устройства вычислительных ресурсов. При запуске на данной платформе используемой сверточной нейронной сети Tiny-YOLOv2 скорость обнаружения объектов на видеозаписи с разрешением 1280×720 составила приблизительно 5 кадров в секунду.

Безбородов А.К. (автор)

Павлов В.А. (научный руководитель)