

СРАВНЕНИЕ МЕТОДОВ ОТСЛЕЖИВАНИЯ ОБЪЕКТОВ В РЕАЛЬНОМ ВРЕМЕНИ

магис. Жданеня О. А., доц. к.т.н. Шемаров А. И.

Белорусский государственный университет информатики и радиоэлектроники

При создании систем компьютерного зрения задача отслеживания объектов имеет большое значение. Она имеет множество вариантов прикладного применения. Для её решения существует множество методов, однако задача не теряет своей актуальности. Каждый год появляются новые методы позволяющие повышать эффективность решения поставленной задачи. Особый интерес вызывают методы способные обработать данные в реальном времени. В статье рассматриваются некоторые методы отслеживания объектов, которые появились относительно недавно и позволяющие получить результаты приемлемые для создания современных систем компьютерного зрения. На основе результатов, полученных с MOT Challenge, производится их сравнение с алгоритмом DeepSORT, который часто используется при решении задач отслеживания объектов.

1. Введение

Отслеживание объектов одно из наиболее востребованных направлений в компьютерном зрении. Процесс отслеживания объектов обычно состоит из трёх этапов: обнаружение объектов, вычисление метрик схожести и сопоставление объектов. В отслеживании объектов очень важно обеспечить получение обнаружений объектов с высокой степенью достоверности, разрешение ситуаций с потерей обнаруженного объекта, а также выходом обнаруженных объектов из кадра. На результаты отслеживания оказывает влияние перекрытие объектов другими объектами.

Задача отслеживания объектов широко применяется для решения практических задач. Решение задачи заключается в построении траекторий для выделенных на видео объектов. Удачное решение задачи может расширить варианты практического применения, например, для интеллектуального анализа видео, взаимодействия человека с компьютером, распознавания движений. Большинство существующих методов решают задачу в два отдельных этапа. Первый этап – этап обнаружения объектов на каждом кадре, для чего используются детекторы. Второй этап – связывание обнаруженных объектов и имеющихся путей. Для этого, сначала, извлекаются признаки повторной идентификации для каждого обнаруженного объекта. Затем, происходит связывание обнаруженных объектов и существующих путей по извлечённым признакам, при необходимости создаются новые пути или удаляются старые.

2. Методы отслеживания

В настоящее время существует множество методов отслеживания объектов. Они имеют свои особенности работы, что может накладывать ограничения на их применение для каждой конкретной реализации системы.

Наиболее популярным методом является DeepSORT[1]. В различных источниках часто встречается упоминание этого метода как наилучшего и наиболее точного. Однако с момента появления этого метода прошло достаточно много времени и появились новые методы, позволяющие решать задачу не хуже, а иногда и лучше чем DeepSORT по комплексу параметров.

Рассмотрим некоторые методы отслеживания объектов, в том числе и DeepSORT, как наиболее популярный метод, с целью получения сравнительных характеристик для анализа.

Выбор методов отслеживания объектов производился на основе нескольких факторов. Методы были проверены на MOT Challenge[2]. Методы могут быть применены для отслеживания объектов в реальном времени. Отобранные методы имеют открытый исходный код или опубликованное подробное описание метода, которое позволяет воспроизвести метод. А также, отобранные методы имеют высокую эффективность, которая оценивалась с помощью метрик MOTA[3] и ID F1[4].

2.1 DeepSORT

DeepSORT основан на алгоритме Simple online and real-time tracking[5]. Для улучшения алгоритма, была добавлена информация о внешности объекта для сравнения обнаруженных объектов на разных кадрах и сопоставления их друг с другом. Это позволяет отслеживать объекты через более долгий период их перекрытия или выхода из кадра с последующим возвращением, и эффективно уменьшить количество смен идентификаторов у объектов.

SORT основан на фильтре Калмана и покадровом связывании данных. Фильтр Калмана применяется для построения путей. Покадровое связывание данных основано на Венгерском методе и использовании ассоциативной метрики для сравнения обнаруженных объектов, которая измеряет пересечение ограничительных рамок.

Алгоритм SORT имеет проблемы с отслеживанием при перекрытии объектов или их исчезновении из кадра с последующим возвращением. Для преодоления этой проблемы ассоциативная метрика была заменена на более информативную метрику, основанную на внешнем виде объекта. Для этого используется свёрточная нейронная сеть, обученная различать пешеходов. Для обучения нейронной сети использовался набор данных MARS[6] для повторной идентификации людей. С помощью интеграции этой нейронной сети увеличивается устойчивость метода к пропускам и перекрытиям объектов, при этом сохраняется простота реализации метода, его эффективность и возможность применения для решения задач в реальном времени.

2.2 ByteTrack

BYTE[7] простой и эффективный метод связывания данных для отслеживания объектов. Отличительной особенностью данного метода является обработка обнаруженных объектов, как с высокой вероятностью обнаружения объекта, так и с низкой. Схема работы метода BYTE представлена на рисунке 1. Сначала, связываются объекты с высокой вероятностью обнаружения и имеющиеся пути. После выполнения данного шага некоторые пути остаются без связей, так как они не совпадают ни с одним обнаруженным объектом с высокой вероятностью. Такое обычно бывает в случае перекрытия объектов, их размыва при движении или изменении размера. В таком случае происходит связывание обнаруженных объектов с низкой вероятностью обнаружения и оставшихся путей.

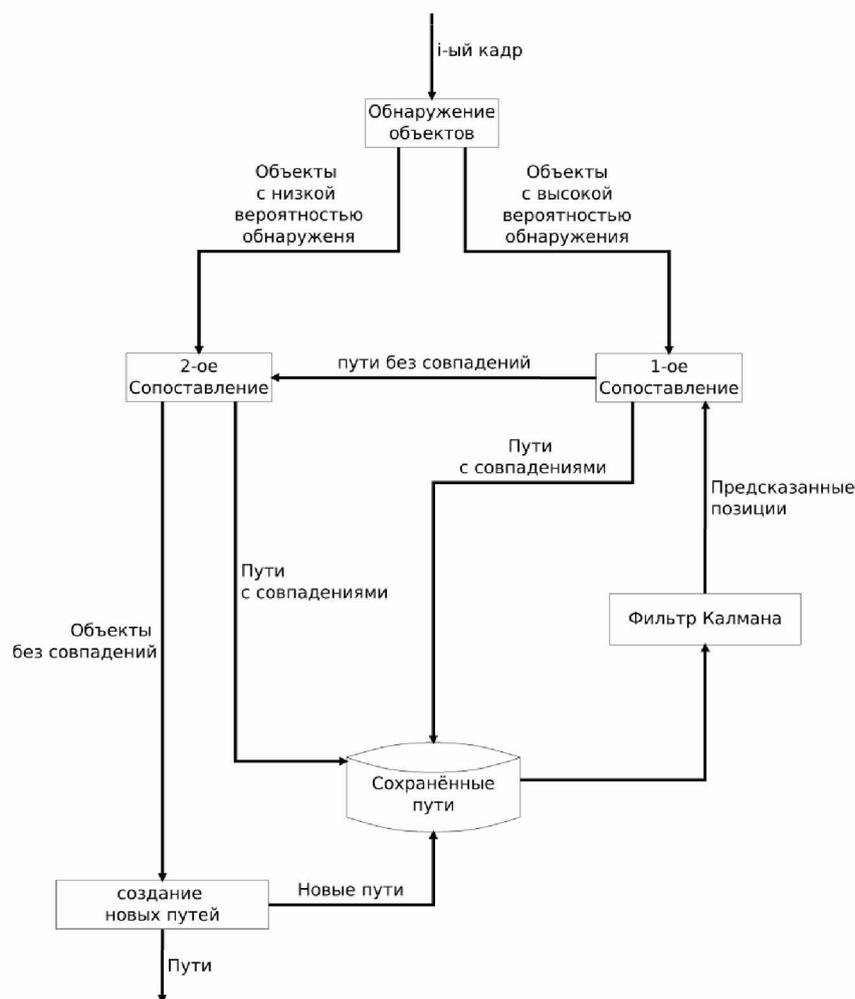


Рисунок 1 – Схема работы метода BYTE

Для получения приемлемого результата при отслеживании нескольких объектов был создан алгоритм ByteTrack, полученный путём использования высокопроизводительного детектора YOLOX[8] в сочетании с методом BYTE.

2.3 SportsTrack

SportsTrack[9] разрабатывался в первую очередь для отслеживания спортсменов во время игр, таких как баскетбол или футбол, или при проведении забегов. Чаще всего методы отслеживания применяют для обработки видео со статичных камер, у которых фон не изменяется, и люди обычно не попадают в последовательность по несколько раз за относительно короткий период.

В спортивных сценах приходится сталкиваться с такими проблемами как размытие объектов из-за движения и перекрытие объектов на кадрах, что ведёт к получению низкой вероятности обнаружения объектов, с неправильными результатами обнаружений объектов с высокой вероятностью. Некоторые методы отслеживания используют двухэтапную стратегию связывания данных с высокой и низкой вероятностью обнаружения. Они устанавливают низкий приоритет для обнаружений объектов с низкой вероятностью. Такой тип стратегии влияет на связывание обнаружений объектов с более высокой вероятностью. Поэтому в развитие предлагается метод Sports matching.

Sports Matching – это трёхэтапный метод связывания данных. Первый этап – связывание всех обнаружений и путей с заданным жёстким пороговым значением. Второй этап – связывание обнаружений с высокой оценкой и оставшихся путей. Последний этап – связывание обнаружений с низкой оценкой и оставшихся путей. На рисунке 2 представлена схема работы метода.

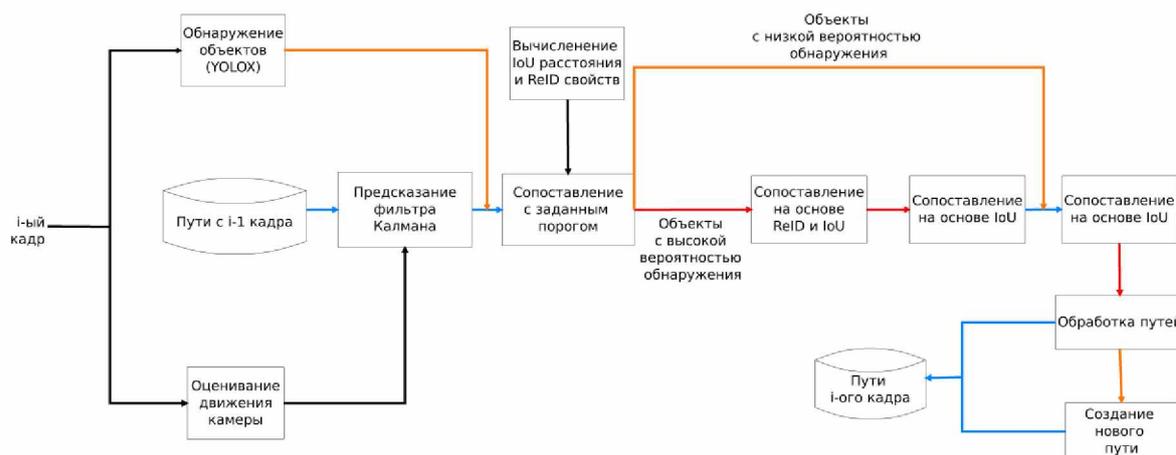


Рисунок 2 – Схема работы SportsTrack

2.4 Strong-TransCenter (STC)

Strong-TransCenter[10] основан на методе TransCenter[11]. Основная часть TransCenter состоящая из использования энкодера-декодера осталась без изменений. На вход поступают два последовательных кадра, при этом на выходе получается тепловая карта с расположением объектов. Изменения в этом алгоритме связаны с изменением обработки путей и связыванием путей и обнаруженных объектов.

Блок связывания объектов и путей выполняет обработку в два этапа, зависящих от вероятности обнаруженных объектов, как в алгоритме ByteTrack, и введением третьего этапа связанного с повторной идентификацией. На первом этапе происходит связывание активных путей и обнаруженных объектов, имеющих высокую вероятность обнаружения. На втором этапе – связываются оставшиеся активные пути с объектами, имеющими низкую оценку обнаружения. На третьем этапе происходит связывание оставшихся обнаруженных объектов и неактивных путей, и предпринимается попытка восстановить пути, т.е. вернуть их в состояние активных.

Все три этапа были адаптированы так, чтобы включить в себя оценку близости объектов, а именно GIoU, и оценку схожести внешности, основанную на нейронной сети FastReID.

На рисунке 3 представлена схема работы метода.

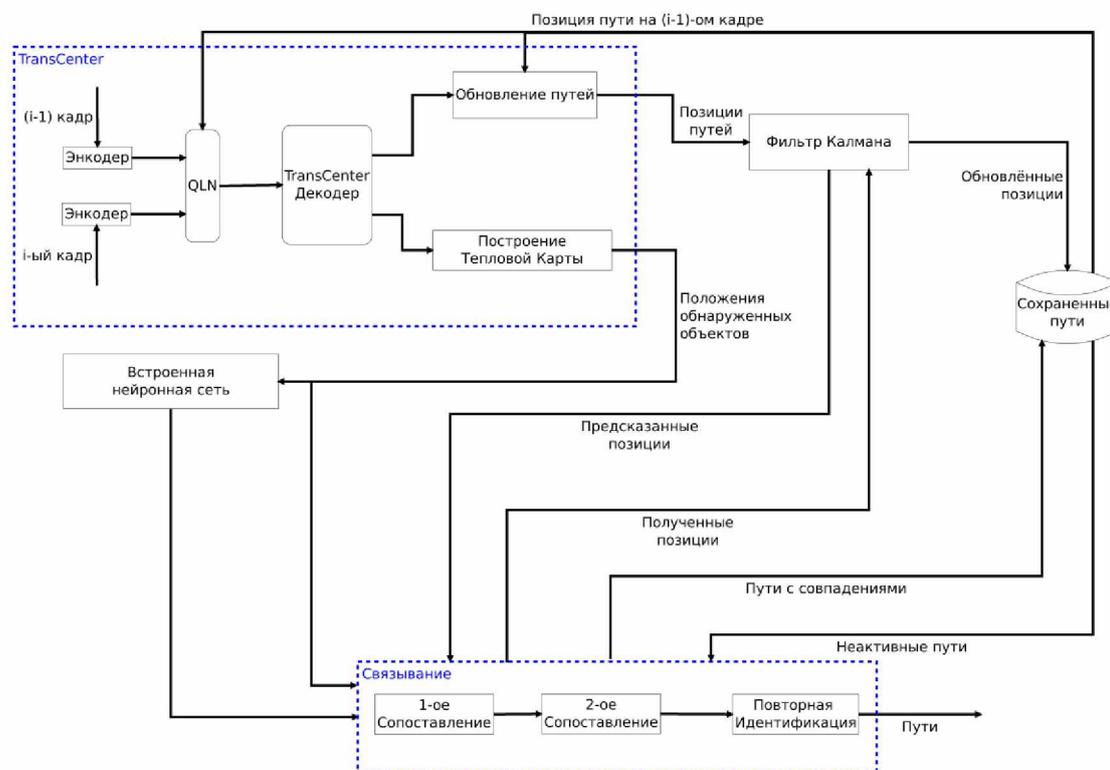


Рисунок 3 – Схема работы Strong-TransCenter

3. Сравнение методов

Для сравнения методов использовались результаты, опубликованные на MOT Challenge. Для рассматриваемых методов отслеживания объектов использовались результаты измерений метрик с тестовых наборов данных MOT16[12] и MOT20[13]. Для DeepSORT доступны данные только для MOT16. Для SportsTrack, Strong-TransCenter и ByteTrack использовались данные с MOT20.

Для сравнения используются метрики MOTA и ID F1. В таблице 1 приведены полученные значения метрик для методов, полученных на разных наборах данных, а также год прохождения проверки на тестовых наборах данных.

Таблица 1. Сравнение методов отслеживания объектов

Метод	MOTA, %	ID F1, %	Год публикации результатов
SportsTrack	74.5 (MOT20)	73.7 (MOT20)	2022
STC Tracker	73 (MOT20)	67.6 (MOT20)	2022
ByteTrack	67.0 (MOT20)	70.2 (MOT20)	2022
DeepSORT	61.4 (MOT16)	62.2 (MOT16)	2017

Вывод

В работе были рассмотрены три метода отслеживания объектов, которые имеют одни из лучших показателей метрик MOTA и ID F1 среди множества предлагаемых методов. Согласно приведённым измеренным метрикам они превосходят по своей эффективности и точности метод DeepSORT, который на данный момент является наиболее популярным и часто используемым методом. Однако новые методы требуют выполнения большего количества вычислений. Это можно отнести к незначительным недостаткам анализируемых методов, так как с развитием аппаратных платформ обеспечивающим увеличение вычислительных мощностей и ростом их доступности, новые методы смогут применяться наряду с методом DeepSORT, а в ряде случаев смогут и заменить его.

Литература

1. N.Wojke, A.Brewley, D.Paulus. Simple online and realtime tracking with a deep association metric // Preprint arXiv:1703.07402v1, 2017.
2. P.Dendorfer, A.Osep, A.Milan, K.Schindler, D.Cremers, I.Reid, S.Roth, L.Leal-Taixe. MOTChallenge: A Benchmark for Single-Camera MultipleTarget Tracking // Preprint arXiv:2010.07548v2, 2020.
3. K.Bernardin, R.Stiefelhagen. Evaluating Multiple Object Tracking Performance: The CLEAR MOT Metrics // EURASIP Journal on Image and Video Processing, 2008. pp. 1–10.
4. Ristani, E., Solera, F., Zou, R., Cucchiara, R. & Tomasi, C. Performance Measures and a Data Set for Multi-Target, Multi-Camera Tracking // European conference on Computer Vision on Benchmarking Multi-Target Tracking, 2016. pp. 1–18.
5. Bewley, G. Zongyuan, F. Ramos, and B. Uppcroft, Simple online and realtime tracking // IEEE International Conference on Image Processing, 2016. pp. 3464–3468.
6. L. Zheng, Z. Bie, Y. Sun, J. Wang, C. Su, S. Wang, and Q. Tian. MARS: A video benchmark for large-scale person re-identification // European Conference on Computer Vision, 2016. pp. 868–884.
7. Y. Zhang, P.Sun, Y.Jiang, D.Yu, F.Weng, Z.Yuan, P.Luo, W.Liu, X.Wang. ByteTrack: Multi-Object Tracking by Associating Every Detection Box // Preprint arXiv:2110.06864v3, 2022.
8. Z. Ge, S. Liu, F. Wang, Z. Li, and J. Sun. Yolox: Exceeding yolo series // Preprint arXiv:2107.08430, 2021.
9. J.Wang, Y.Peng, X.Yang, T.Wang, Y.Zhang. SportsTrack: An Innovative Method for Tracking Athletes in Sports Scenes // Preprint arXiv:2211.07173v1, 2022.
10. A.Galor, R.Orfaig, B. Bobrovsky. Strong-TransCenter: Improved Multi-Object Tracking based on Transformers with Dense Representations // Preprint arXiv:2210.13570v1, 2022.
11. Y.Xu, Y.Ban, G.Delorme, C.Gan, D.Rus, X.Alamenda-Pineda. TransCenter: Transformers with Dense Queries for Multiple-Object Tracking // Preprint arXiv:2103.15145v1, 2021.
12. A.Milan, L.Leal-Taixe, I.Reid, S.Roth, K.Schindler MOT16: Benchmark for Multi-Object Tracking // Preprint arXiv:1603.00831v2, 2016.
13. P.Dendorfer, H. Rezatofighi, A.Milan, J.Shi, D.Cremers, I.Reid, S.Roth, K.Schindler, L.Leal-Taixe MOT20: A benchmark for multi object tracking in crowded scenes // Preprint arXiv:2003.09003v1, 2020.

COMPARISON OF REAL TIME OBJECT TRACKING METHODS

ms. stud. Zhdanenia O.A., ass. prof., PhD Shemarov A.I.

Belarusian State University of Informatics and Radioelectronics

Object tracking task is important part of creating computer vision systems. It has many applications. There are many methods to solve the task. But it doesn't lose its relevance. Almost every year new methods that supposed to increase efficiency of solving the task come out. Of particular interest are methods that can track objects in real time. The article discusses some methods for object tracking that have appeared relatively recently and allow obtaining results acceptable for creating modern computer vision systems. Based on results obtained from MOT Challenge they are compared with the DeepSORT algorithm which is often used for object tracking.