

# Применение алгоритмов машинного обучения для поиска рельсовой колеи

Н. М. Гаврилова<sup>1</sup>, И. А. Дейлид<sup>2</sup>, С. А. Молодяков<sup>3</sup>

Санкт-Петербургский политехнический университет

Петра Великого

<sup>1</sup>spbstung@yandex.ru, <sup>2</sup>ivanov.dale@gmail.com,

<sup>3</sup>molod@ics2.ecd.spbstu.ru

Е. О. Болтенкова<sup>4</sup>, Д. А. Никитин<sup>5</sup>, П. А. Попов<sup>6</sup>

Санкт-Петербургский филиал НИИАС

<sup>4</sup>kitkat52@yandex.ru, <sup>5</sup>d.nikitin@vniias.ru,

<sup>6</sup>p.popov@vniias.ru

**Аннотация.** Рассматривается задача определения рельсовой колеи на пути движения поезда по видеоизображению с камеры, установленной на локомотиве. Предлагается метод, основанный на комбинировании нескольких алгоритмов машинного обучения. Рассматривается поэтапное применение алгоритмов поиска признаков объектов в кадре, их комплексная классификация, проверка на достоверность, построение графа для создания маски пути. Приводятся промежуточные и итоговые примеры применения алгоритмов на видеоизображениях железнодорожных путей.

**Ключевые слова:** компьютерное зрение; машинное обучение; SURF; бинарная классификация; метод ближайших соседей; метод опорных векторов; метод k-средних; граф ближайших соседей; преобразование перспективы

## I. ВВЕДЕНИЕ

В настоящее время области применения компьютерного зрения непрерывно расширяются, разрабатываются все новые алгоритмы обработки и распознавания изображений. Для лучшего решения задач компьютерного зрения создаются видеокамеры повышенного разрешения, используются быстродействующие многоядерные видеопроцессоры, применяются лидары, радары и другие датчики информации об окружающей среде. Комплексирование данных, полученных с разных датчиков, позволяет создавать системы с надежным распознаванием объектов, которые можно использовать в разных погодных условиях.

Транспорт является той областью, в которой применение компьютерного зрения дает ощутимые результаты. Многие современные автомобили уже включают элементы автопилотирования, которые построены на основе обработки видео. Известны успехи по созданию беспилотного автомобиля компании Tesla [1], отвечающего третьему уровню системы автопилотирования сообщества автомобильных инженеров (SAE) [2]. Проводится активная работа по внедрению компьютерного зрения на железнодорожном транспорте. Применение компьютерного зрения позволяет сделать существенный шаг в построении автопилотируемых поездов, обеспечении повышенной безопасности движения, локализации местоположения поезда и в целом

повысить автоматизацию железнодорожных перевозок. Так проведена большая исследовательская работа в университете Зальцбурга (University of Salzburg, Austria), в рамках которой был оборудован датчиками экспериментальный локомотив и разработан ряд алгоритмов компьютерного зрения [3].

В научно-исследовательском и проектно-конструкторском институте информатизации, автоматизации и связи на железнодорожном транспорте (компания «НИИАС») также проводятся работы по автоматизации движения железнодорожного транспорта, по созданию автономных поездов, применению систем компьютерного зрения. В настоящее время один из маневровых локомотивов оборудован двумя камерами ближней и дальней областей наблюдения, лидаром и радаром (рис. 1). Данные с камер в режиме реального времени могут поступать в лабораторию Санкт-Петербургского филиала НИИАС.



Рис. 1. Фотография локомотива, оборудованного датчиками системы технического зрения



Рис. 2. Оригинальное изображение железнодорожных путей

На рис. 2 представлен оригинальный кадр изображения, полученный с камеры разрешения 2К. Одной из главных задач обработки изображения является определение рельсовой колеи на пути движения локомотива с помощью видеокамеры. Причем желательно определить одну рельсовую колею, по которой будет двигаться локомотив.

Основные трудности решения задачи связаны с наличием стрелок, с необходимостью работы процессора видеосистемы в реальном времени и с необходимостью учитывать влияние различных погодных явлений (снег, дождь).

Для определения рельсовой колеи можно использовать ряд алгоритмов. Один из них – использование методов на основе сегментационных нейронных сетей (SegNet, Unet, FCN и др.) [4]. Основные недостатки этих методов в больших вычислительных затратах на кадр, а также трудностями в интерпретации результатов работы глубокой нейронной сети, из-за чего совместно с этим алгоритмом разумно использовать ещё несколько, чтобы понизить вероятность ошибки принятия неверного решения.

В данной статье будет рассмотрен алгоритм определения рельсовой колеи на пути движения с помощью анализа признаков объектов (англ. *features*), будет поэтапно показана работа алгоритма в задачах классификации, удаления шумов и построения из отдельных *features* маски рельсовой колеи. Представляемый алгоритм можно рассматривать как один из алгоритмов сегментирования изображения.

## II. РАЗРАБОТАННЫЙ АЛГОРИТМ

### A. Определение ключевых точек на изображении

Первый этап алгоритма – извлечение и классификация признаков из изображения. Для этого используется метод *speeded up robust features (SURF)*, который является быстрым методом обнаружения устойчивых признаков изображения [5]. На выходе SURF-этапа получаем информацию в виде координат ключевых точек и их описания (дескрипторы). Эта информация в дальнейшем используется для классификации точек.



Рис. 3. Пример позитивного (а) и негативного (б) изображений

Предварительно был выделен набор данных, состоящий из изображений рельсовой колеи – позитивный пример и изображений фона – негативный пример (рис. 3). В ходе тренировки из каждого изображения извлекаются найденные признаки, а их дескрипторы помечаются меткой класса (в зависимости от того, к какому набору данных относится изображение).

### B. Классификация ключевых точек

На втором этапе проводится классификация выделенных ключевых точек. В работе была использована комбинация из двух достаточно простых и популярных методов классификации: метод *k*-ближайших соседей (*K-nearest neighbors, KNN*) [6] и метод опорных векторов (*support vector machines, SVM*) [7]. Стоит отметить особенности тренировки данных классификаторов: для SVM'a рекомендуется подбирать тюнинг-параметры таким образом, чтобы избежать переобучения, поскольку признаки, найденные в одном наборе данных, могут с достаточно большой вероятностью оказаться и в другом. Для метода ближайших соседей оптимизация несколько сложнее: после создания обучающей выборки над ней производится кластеризация алгоритмом *K-Means* [8], после чего метка каждого центроида определяется подавляющим количеством признаков (позитивных или негативных). Затем создаётся новая обучающая выборка для алгоритма, состоящая из меньшего количества примеров.

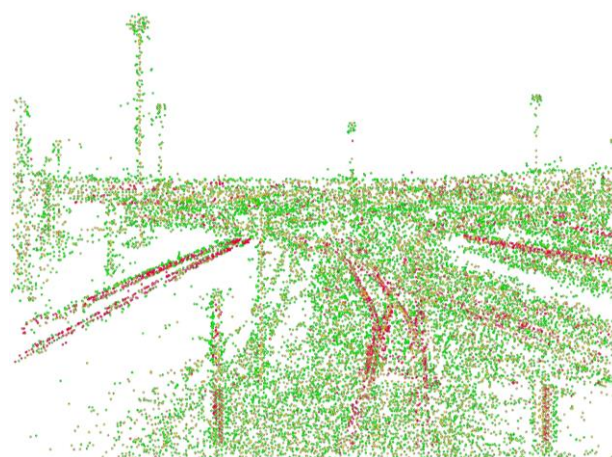


Рис. 4. Результат классификации ключевых точек

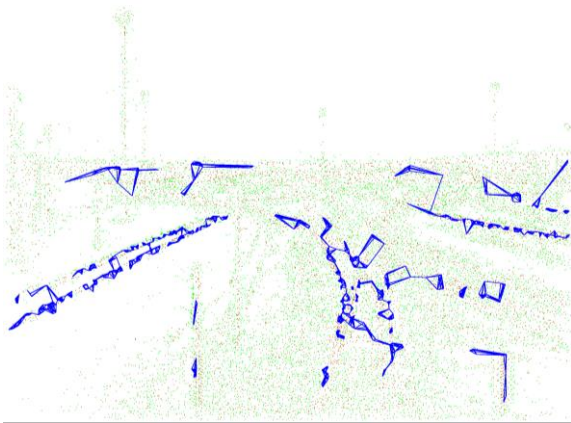


Рис. 5. Граф рельсовой колеи, составленный из ключевых точек с числом соседей равным 6

Два главных преимущества такого подхода состоят в том, что значительно уменьшается количество тренировочных данных для метода ближайших соседей (в данном примере 500 вместо 7000), что увеличит скорость работы классификатора, во-вторых таким образом можно избежать переобучения, а в-третьих упрощается задача выбора параметра количества соседей. Теперь классифицируются все точки на экране, в случае, если оба классификатора дали положительный результат – данная точка считается принадлежащей рельсовой колее.

На рис. 4 представлен результат комплексирования данных классификаторов на основе метода опорных векторов и метода k-ближайших соседей: красные точки – рельсы, определенные двумя классификаторами; желтые – точки, определенные как рельсы лишь одним из классификаторов; зеленые – фон.

#### С. Удаление шумов

Однако, любой алгоритм машинного обучения не может гарантировать абсолютно верного результата, поэтому следующим этапом определения рельсовой колеи будет удаление шумов или ложных-положительных срабатываний классификатора. Поскольку можно пропустить точки, определенные как не относящиеся к рельсам, но являющиеся ими (достаточно знать некоторые точки контуров рельсов, а остальные аппроксимировать), данные ошибки классификатора будем игнорировать. Гораздо больше проблем могут доставить точки, ложно определенные как относящиеся к рельсам, поскольку на следующем этапе, когда они будут участвовать в составлении маски, возможно получить искажения. Для их удаления воспользуемся свойством компактности. Логично предположить, что большинство точек рельсовой колеи будут находиться рядом друг с другом поскольку на кадре они также должны находиться рядом, поэтому следует убрать все те точки, среди которых ближайшие N соседей определены как не рельсы.



Рис. 6. Кадр изображения с маской, полученной после 15-кратного выполнения морфологического закрытия

#### Д. Составление маски

Последним этапом является составление маски рельсовой колеи. Для этого применяется метод NNG (Nearest neighbor graph) [9]. Составим на изображении граф, по следующему принципу: соединим каждую точку с K ближайших к ней. Таким образом контурами данных фигур покрываются участки изображения, относящиеся к рельсам. Варьируя параметр K, можно избежать искажений и ошибок в построении маски (уменьшая параметр K), либо уменьшить вероятность разрыва участков рельсовой колеи (увеличивая параметр K). На рис. 5 представлен граф рельсовой колеи, составленный из ключевых точек с числом соседей равным 6. Чтобы из полученных графов создать непосредственно маску – многократно выполним морфологическую операцию закрытия над получившимся графом (рис. 6). На рис. 6 видно, что маска накрывает не все железнодорожные пути.

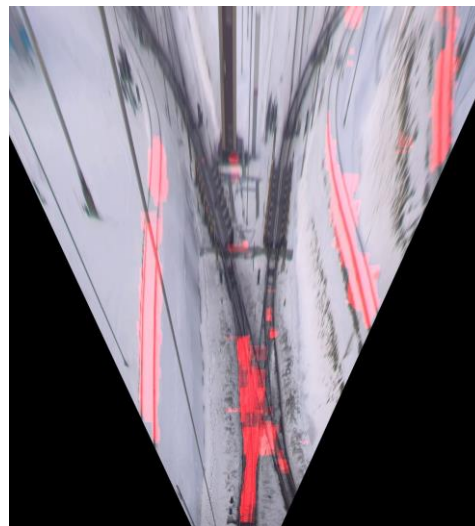


Рис. 7. Кадр с наложенной маской после преобразования перспективы



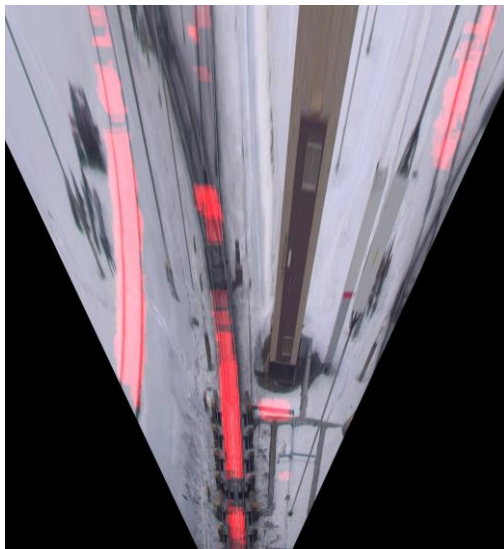


Рис. 8. Изображение рельсовой колеи с наложенной усредненной по последовательным 5 кадрам маской

Не полное покрытие связано с тем, что тренировка SURF-этапа проводилась на летних примерах изображения (рис. 3), а рельсовая колея представлена на зимних видео (рис. 2). Для того, чтобы получить более полное покрытие рельсов маской предлагается накапливать данные полученные от предыдущих кадров, усредняя их маски. Это позволяет более устойчиво определять колею, по которой движется локомотив.

Для практического использования найденной маски пути необходимо сопоставить данные о ней с другими различными датчиками и методами семантической разметки изображения. Поэтому над кадром выполняется обратное преобразование перспективы (Bird View [10] проекция), что в дальнейшем позволяет участвовать полученным результатам в комплексном описании окружающей локомотив среды. На рис. 7 показана карта рельсовой колеи с видом сверху, на которую наложена маска, вычисленная за 1 кадр. На рис. 8 показано изображение рельсовой колеи, на которое наложена маска, усредненная по последовательным 5 кадрам.

### III. РЕЗУЛЬТАТЫ

В результате выполнения ряда экспериментов определена точность классификаторов: 83% для метода ближайших соседей, 64% для метода опорных векторов. Определено, что комбинирование данных методов, а также дополнительных операций над полученным множеством результатов позволили получить приемлемую приближенную маску рельсовой колеи железнодорожного пути.

### IV. ЗАКЛЮЧЕНИЕ

Таким образом, разработан метод поиска рельсовой колеи, который представляет собой комбинацию алгоритмов машинного обучения. Для повышения точности классификатора использована процедура последовательного применения алгоритмов k-ближайших соседей, опорных векторов и покадрового усреднения. Экспериментально показана работоспособность разработанного алгоритма. Метод можно рассматривать как один из способов сегментирования изображения. Основные преимущества алгоритма, реализующего метод, связаны с минимальной подготовкой тренировочной выборки, возможностью осознанно влиять на работу алгоритма и анализировать его работу для дальнейшего улучшения.

Описанный метод сегментирования является элементом решения основных задач технического зрения: позиционирования и определения препятствий на пути движения поезда. Дальнейшее развитие системы связано с использованием комплексирования данных, полученных с лидара и радаров, а также применением альтернативных алгоритмов сегментирования для поиска рельсовой колеи.

### СПИСОК ЛИТЕРАТУРЫ

- [1] Shantanu I., Phute M. Tesla autopilot: semi autonomous driving, anuptick for future autonomy // International Research Journal of Engineering and Technology. 2016. V. 3. Iss. 9. P. 369-372.
- [2] SAE Levels of Driving Automation <http://cyberlaw.stanford.edu/blog/2013/12/sae-levels-driving-automation>
- [3] Gebauer O., Pree W., Stadlmann B. Autonomously Driving Trains on Open Tracks – Concepts, System Architecture and Implementation Aspects // Information Technology. 2012. V. 54, Iss. 6. P. 266-278. DOI: 10.1524/itit.2012.0689
- [4] Girshick R., Donahue J., Darrell T., Malik J. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation // IEEE Conference Computer Vision and Pattern Recognition (CVPR), 2014 DOI: 10.1109/CVPR.2014.81.
- [5] Bay H., Tuytelaars T., Luc Van Gool SURF: Speeded Up Robust Features // Computer Vision and Image Understanding (CVIU). 2008. V. 110, No. 3. P. 346-359. Doi:10.1016/j.cviu.2007.09.014
- [6] Guo G., Wang H., Bell D., Bi Y., Greer K. (2003) KNN Model-Based Approach in Classification. In: Meersman R., Tari Z., Schmidt D.C. (eds) On The Move to Meaningful Internet Systems 2003: CoopIS, DOA, and ODBASE. OTM 2003. Lecture Notes in Computer Science, V. 2888. Springer. doi.org/10.1007/978-3-540-39964-3\_62
- [7] Hsu C.W., Chang C.C., Lin C.J. A practical guide to support vector classification. 2003 <http://mcs.une.edu.au/doc/libsvm/guide.pdf>
- [8] Kanungo T., Mount D.M., Netanyahu N.S., Piatko C., Silverman R., Wu A.Y. An Efficient k-Means Clustering Algorithm: Analysis and Implementation // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2002. V.24. N.7. P.881-892. doi:10.1109/TPAMI.2002.1017616
- [9] David Eppstein Michael S. Paterson Frances F. Yao On Nearest-Neighbor Graphs // Discrete & Computational Geometry. 1997 V. 17. No. 3. P. 263-282. <https://www.ics.uci.edu/~eppstein/pubs/EppPatYao-DCG-97.pdf>
- [10] Liu YC., Lin KY., Chen YS. Bird's-Eye View Vision System for Vehicle Surrounding Monitoring // Robot Vision. 2008. Lecture Notes in Computer Science, V. 4931. doi.org/10.1007/978-3-540-78157-8\_16