

Р.А. Горбачев, М.Н. Зарипов, Д.Л. Шишков

**ИНТЕЛЛЕКТУАЛЬНАЯ СИСТЕМА ТЕХНИЧЕСКОГО ЗРЕНИЯ ДЛЯ
ОБНАРУЖЕНИЯ ПРЕПЯТСТВИЙ И ПРЕДСКАЗАНИЯ ПОВЕДЕНИЯ
ДВИЖУЩИХСЯ ОБЪЕКТОВ НА ЖЕЛЕЗНОДОРОЖНЫХ ПУТЯХ**

В настоящее время повышение качества предоставляемых транспортных и логистических услуг напрямую связано с внедрением новых и модернизацией существующих технологий информатизации и цифровизации. Одной из наиболее актуальных задач, решаемых с помощью внедрения цифровых технологий в существующие технологические процессы, является повышение безопасности движения поездов. Анализ отечественных и зарубежных работ, посвящённых разработке систем повышения безопасности движения поездов показал, что одним из методов решения поставленной задачи является разработка и внедрение систем технического зрения для обнаружения объектов инфраструктуры и препятствий по ходу движения поезда. Это особенно актуально при увеличении скоростей движения поездов, когда машинисту бывает сложно правильно оценить сложившуюся ситуацию и принять оперативное решение. В данной работе описана реализация системы технического зрения для беспилотных поездов. В ее рамках был реализован новый подход к обучению узкоспециализированной масочной нейронной сети. Основной задачей этой системы является распознавание препятствий и фигур человека на фоне железнодорожной инфраструктуры, определения их местоположения относительно рельсовых путей и оценки этой ситуации с точки зрения безопасности движения. Для получения более качественной маски был использован подход одновременного использования изображений стандартных камер CVS и камер с более высоким разрешением. Данный метод способен повысить качество распознавания, особенно на больших расстояниях, когда интересующий объект не заметен в сложной среде окружающей его обстановки. Выполненная работа показала хорошие результаты по идентификации объектов на железнодорожных путях. Создание прототипа такой системы и оснащение ей тягового подвижного состава позволит реализовать своевременное обнаружение препятствий и людей на пути поезда, что способствует повышению уровня безопасности движения поездов.

Система компьютерного зрения; нейронные сети; метод Лукаса–Канаде; карта глубины; фотограмметрия.

R.A. Gorbachev, M.N. Zaripov, D.L. Shishkov

**AN INTELLIGENT SYSTEM OF TECHNICAL VISION FOR DETECTING
OBSTACLES AND PREDICTING THE BEHAVIOR OF MOVING OBJECTS
ON RAILWAY TRACKS**

Currently, the improvement of the quality of transport and logistics services provided is directly related to the introduction of new and modernization of existing technologies of informatization and digitalization. One of the most urgent tasks solved by the introduction of digital technologies into existing technological processes is to improve the safety of train traffic. The analysis of domestic and foreign works devoted to the development of train safety improvement systems has shown that one of the methods of solving the task is the development and implementation of vision systems for detecting infrastructure objects and obstacles in the course of train movement. This is especially true when train speeds increase when it is difficult for the driver to correctly assess the current situation and make an operational decision. This paper describes the implementation of a vision system for unmanned trains. Within its framework, a new approach to the training of a highly specialized mask neural network was implemented. The main task of this system is to recognize obstacles and human figures against the background of the railway infrastructure determine their location relative to the tracks and assess this situation from the point of view of traffic safety. To obtain a higher-quality mask, the approach of simultaneous use of images of standard CVS cameras and cameras with the higher resolution was used. This method is able to

improve the quality of recognition, especially at large distances, when the object of interest is not noticeable in the complex environment surrounding it. The work performed has shown good results in identifying objects on railway tracks. The creation of a prototype of such a system and equipping it with traction rolling stock will allow for the timely detection of obstacles and people on the train path, which contributes to improving the level of train safety.

Computer vision system; neural networks; Lucas–Canada method; depth map; photogrammetry.

1. Введение. В настоящее время из-за увеличения плотности железнодорожных перевозок любые отклонения от установленного графика приводят к значительным потерям качества движения, например, к увеличению задержек поездов или их отмене. Одной из наиболее распространенных предпосылок такого рода является появление препятствия или присутствие человека на железнодорожных путях, которые мешают движению поезда и способны привести не только к задержкам в движении, но и к авариям.

Повышение качества предоставляемых транспортных и логистических услуг напрямую зависит от внедрения новых и улучшения существующих технологий информатизации и цифровизации, особенно на железнодорожном транспорте ввиду постоянно растущего трафика и задачи наиболее эффективного использования уже существующих ресурсов для его обеспечения. Однако разработка и внедрение технологий в данной области подчинено требованиям безопасности движения. Одним из методов повышения безопасности на железной дороге является разработка систем обнаружения посторонних объектов на путях, а также их анализ [1]. Поэтому разработка систем идентификации и классификации объектов на путях является одной из важнейших задач в системах управления железнодорожным движением. В данной области особенно хорошо себя показали технологии компьютерного зрения и искусственного интеллекта. Такие системы не только обнаруживают препятствия, но и классифицируют их для определения последующих действий, например, возможности дальнейшего движения или аварийной остановки поезда.

Развитие представленной в данной работе технологии способствует цифровизации железнодорожного транспорта [2] и особенно востребовано для использования в движении беспилотных поездов, так как система технического зрения выполняет ключевую роль в оценке текущей дорожной ситуации.

2. Обзор. Задачам распознавания для железных дорог посвящен ряд работ [3–5]. Рассмотрим основные идеи, представленные в них.

В работе [4] предложен метод обнаружения препятствий путем сравнения входного и эталонного изображений с камеры лобового обзора поезда. Одним из новшеств данной работы является обнаружение неизвестных классов объектов, т.е. которых не было в обучающей выборке с помощью предложенного метода вычитания фона, который можно применять и к нестационарным камерам.

В работе [5] представлен прототип устройства TRINETRA, который обещает легкое движение поездов в неизвестных условиях, таких как условия нулевой видимости, туман, смог и сильный дождь. Прототип основан на интеграции камеры, системы радиобнаружения и определения дальности (RADAR) и усиления инфракрасного (ИК) света за счет стимулированного излучения (LASER). Комбинированный короткий, средний, и система радарных датчиков дальнего действия используется для непрерывного обнаружения препятствий в слепых зонах локомотивов на трассе, особенно для помощи в предотвращении столкновений на высокой скорости. Экспериментальные результаты испытания прототипа для расстояний от 2 м до 2 км на движущемся поезде показывают, что разработанный прототип эффективно отслеживает препятствия в условиях тумана и смога.

В работе [5] приведена реализация системы обнаружения препятствий, основанной на глубоком обучении нейронной сети MaskR-CNN, как описано в этой статье. Система обнаружения использует модель Mask-RCNN с ResNet101 в качестве своей базовой сети извлечения признаков. Результаты испытаний показали, что точность модели Mask-RCNN с ResNet101 в качестве ее базовой сети извлечения признаков достигла 95,7% и что для этого потребовалось среднее время 0,18 с.

Однако подходы, приведенные в работах [3–5] предназначены только для обнаружения препятствий. Подход, предложенный в данной работе, позволяет не только выделять объекты, но и рассчитывать расстояние до них, а также оценить поведение этих объектов.

3. Постановка задачи. Одной из важнейших проблем, требующих решения на железной дороге является возникновение аварий, вызванных препятствиями. Для этого разрабатываются системы обнаружения препятствий на железнодорожных перегонах [6]. Задачами разработки и развития систем технического зрения для поездов и локомотивов является разработка алгоритмов по анализу цветового пространства, определению местоположения мешающих движению объектов и их идентификации на цифровом изображении. Для решения поставленных задач была реализована система анализа исправности рельсового пути и наличия препятствий на основе анализа плоского (нестереоскопического) изображения. Ключевым компонентом данного подхода является интеллектуальная система технического зрения (Далее - СТЗ). Главной задачей, решаемой с помощью данной СТЗ для беспилотных поездов, является обнаружение и классификация препятствий, возникающих в процессе железнодорожного движения. Распознавание препятствий (люди, повреждения конструкций и т.п.) осуществляется на расстоянии 200–500+ м. Данные объекты в дальнейшем будем называть объектами интереса.

4. Оптимизация входных изображений. Однако при решении задач обнаружения и распознавания объектов, в которых используется автоматическая подстройка параметров видеокамер, очень часто возникают такие проблемы как засветка изображений, затененность и размытость изображений и т.п. (рис. 1).



Рис. 1. Пример неудачных снимков фронтальных камер (засветка)

В связи с этим были проведены дополнительные исследования оптимизации изображений в случаях возникновения таких дефектов и способов их устранения, например, увеличение контрастности и т.п. Для этого для тестового участка железнодорожной трассы на основе видеоданных фронтальных камер была построена 3d-карта объектов в ближней зоне железнодорожных путей, в которую входят такие объекты как столбики, стрелки и другие элементы железнодорожной инфраструктуры, которые визуально могут быть похожи на большом расстоянии на фигуру человека. При необходимости, такая съемка с помощью камер СТЗ дополняется или даже заменяется съемкой с помощью специализированных камер высоко-

го разрешения. Привязка к месту, т.е. точное позиционирование, может осуществляться с помощью высокоточных приемников GPS/ГЛОНАСС в режиме кинематики реального времени и/или с помощью другой системы позиционирования.

Для нейронной сети была реализована система автоматизированного построения дата-сетов обучающих и контрольных выборок объектов интереса из видеофайлов. Для этого осуществляется генерация изображений с объектами интереса, получаемых из реальных видеопотоков, поступающих от видеокамер.

Для проверки системы были проведены исследовательские испытания стенда СТЗ на неиспользуемых участках железнодорожных путей. В процессе этих испытаний были проверены технологии калибровки камер, распознавания объектов интереса на различных расстояниях, в разных условиях освещенности и различных погодных условиях. Создание цифровой модели тестового участка было реализовано с помощью лазерного сканера и серий фотографий с последующей их обработкой посредством алгоритмов фотограмметрии.

СТЗ в данной системе основано на комплексном применении оптических камер и искусственного интеллекта для сопоставления данных съемки с накопленной базой снимков и трехмерных моделей. При этом для корректной и более точной работы совершенно необходима гибкая настройка параметров работы камер СТЗ в зависимости от места, времени, условий освещения и т.п. Фактически настройки камер привязаны к маршруту и времени перемещения подвижного железнодорожного состава.

На подобного рода снимках прежде всего можно обнаруживается наличие резких перепадов освещения в подлежащей контролю окрестности пути следования поезда. В дополнение к этому существует и необходимость одновременного контроля в реальном времени как засвеченных, так и затемненных участков вблизи пути движения поезда. То есть при движении, например, в туннеле нужно одновременно различать хорошо освещенные объекты на выходе из туннеля и затемненные участки вблизи стен туннеля. Другим примером такой ситуации служит движение по мосту, где также необходимо одновременно различать освещенные объекты между сооружениями моста и объекты в затемненных участках сооружения моста.

Все это приводит к тому, что необходимо использовать несколько стереокамер, каждая из которых имеет свои настройки освещенности, то есть контролирует свой диапазон яркости объектов в поле зрения). Получаемые от камер данные будут интегрироваться в единое 3d-облако точек с указанием для каждой точки стереокамеры-источника для предотвращения потерь видеоданных и правильного программного совмещения освещенности и цвета точек облака, полученных от разных видеокамер.

Полученное облако точек далее будет использоваться в 2-х модификациях для 2-х видов обработки.

Плотное точное облако точек будет использоваться для «медленной» фоновой обработки с целью регулярного профилактического визуального контроля оборудования и сооружений станций и путевых сооружений, обнаружения на путях посторонних предметов посредством системы технического зрения поезда. «Медленная» обработка может выполняться на стационарных компьютерах. При этом данные должны передаваться в фоновом режиме с поезда на стационарный компьютер вне зависимости, находится ли поезд на станции или следует по маршруту. Это необходимо для определения необходимости направления ремонтных бригад, а также срочности направления ремонтных бригад (в течение суток, в течение недели и т.д.).

Для «быстрой» обработки данных в реальном времени и для рационального использования вычислительных ресурсов в данной системе применена кусочно-разреженная версия облака точек. Задача быстрой обработки – определение наруше-

ний, требующих немедленной реакции и остановки поезда. Поэтому в ближней окрестности рельсовых путей плотность облака точек должна быть достаточной для выявления явных нарушений в структуре, в составе которой входят рельсы, шпалы, насыпь, а также в рельсовой части стрелочных сооружений. Вне окрестности рельсовых путей плотность облака точек должна быть достаточной для обнаружения сравнимых размером с человеком посторонних объектов, которые отсутствуют в исходном априорном описании путевых и станционных сооружений.

Технологически «быстрая» обработка видеоинформации будет выполняться следующим образом.

Статические кадры ближней окрестности рельсового пути в надлежащем ракурсе будут подаваться на вход искусственно минимизированной нейронной сети *ssdmobilenet*, обученной на распознавание минимальных регулярных участков рельсовых путей и рельсового механизма вблизи стрелок с целью анализа их целостности и соответствия «правильному» образцу. Нейронная сеть реализована и обучена для выявления и локализации таких участков, а также определения их целостности как соответствия критериям, заложенным при обучении.

Кадры дальней окрестности пути анализируются помощью метода Лукаса-Канаде [7], после чего полученное поле скоростей фильтруется для очистки от:

- ◆ движений, обусловленных движением поезда вдоль путей;
- ◆ случайных движений, обусловленных неидеальностью пути и ходовой части;
- ◆ помех посредством ценза по скорости.

Далее эти кадры сравниваются со сделанной ранее точечной картой сооружений, динамически медленно изменяемой согласно видеоданным поездов. Данное сравнение обеспечивает выявление некартированных объектов, линейные размеры которых по вертикали и горизонтали близки к размерам человека. Такие объекты с высокой вероятностью могут оказаться людьми, находящимися в окрестности пути, что в свою очередь потребует остановки поезда для предотвращения столкновения с ними.

«Медленная» фоновая обработка выполняется путем сравнения изображения с «предыдущими» и с «правильными» (то есть сделанными немедленно после выполненного монтажа и верифицированными специалистом-экспертом либо исполнителем ремонта на исправность изображенного оборудования) изображениями посредством специально обученных для этой цели нейронных сетей с целью определения факта наличия и вида неисправности, диагностируемой по отличию фактического изображения от «правильного».

Также при «медленной» обработке определяются объекты, не предусмотренные конструктивно (упавшее дерево, стихийная свалка мусора, и т.д.). Сведения о них будут передаваться эксперту, выносящему решение либо о ликвидации объекта и о ее сроках и приоритетности, либо об игнорировании объекта с занесением его на 3d-карту априорных объектов «быстрой» обработки как безвредного объекта.

В условиях более сложной фоно-целевой обстановки для распознавания препятствий, необходима информация о маршруте поезда. Это позволяет уменьшить рабочую область распознавания и увеличить скорость обработки визуальной информации.

В результате тестирования системы на объектах в железнодорожной инфраструктуре были дополнительно реализованы следующие функции:

- ◆ устранение ложных срабатываний от таких объектов, как, например, различный бытовой мусор;
- ◆ работа на расстояниях от 200м;
- ◆ специализация настроек камер;
- ◆ использование стереопар с большим плечом
- ◆ и другие доработки.

В условиях быстро меняющегося освещения (например, облачная погода с непостоянным присутствием солнца) неотъемлемой частью системы является использование двух-трех комплектов камер с разными настройками или гибкая настройка параметров объективов. Это позволяет повысить эффективность и качество работы системы в переменных погодных условиях.

Совместная работа СТЗ и стационарных камер, передающих информацию о состоянии пути по радиоканалу на приближающийся поезд.

5. Описание реализованной технологии. Система распознавания состоит из следующих аппаратных компонентов:

- ◆ Стереопары с разными объективами, а также стереопара в тепловом диапазоне.

- ◆ Сборка СТЗ с бортовым вычислителем (компьютер, возможно специализированный, с соответствующим ПО).

Для решения поставленной задачи были реализованы следующие программные компоненты системы:

- ◆ Сегментация изображений, полученных с помощью камер. Суть этого подхода заключается в сужении области применения алгоритмов распознавания на область интереса. Данный компонент реализован посредством нейронной сети, выполняющей задачу «Semantic segmentation» для рельсового пути и возможных препятствий с целью выделения их маской. Подробно технология семантической сегментации описана в работах [8, 9]. В данном случае областями интереса являются:

- Непосредственно пути;
- Прилегающие к путям участки и сооружения;
- Области, появление в которых объектов интереса может представлять потенциальную опасность для железнодорожного движения или жизни людей.

- ◆ Стереорекострукция, соединенная с получением данных от нейронной сети об обнаруженных объектах интереса. Данная технология работает с видеоданными и картами глубины и определяет 3d-координаты объектов, а также расстояния до них. В дополнение к этому строится прогноз траектории движения объектов, например, людей. Данный компонент на основе изображения от видеокамеры и полученных от нейронной сети в виде бинарных изображений масок объектов реализует выполнение следующих функций:

а) Проверяет наличие и топологическую целостность видимого рельсового пути. Если рельсовых путей несколько – проверяется тот, который идет от середины нижнего края изображения (то есть тот путь, на котором находится данный поезд). Рельсы должны быть топологически непрерывны, не иметь видимых дефектов и участков с резкими границами, сами рельсы и их ближайшая окрестность не должны быть сверху ничем закрыты.

б) В случае обнаружения людей проверяется, насколько они удалены от железнодорожных путей, является ли расстояние безопасным. Предполагается расстояние обнаружение людей 200 - 500+ метров, что вследствие удаленности сильно снижает точность стереорекострукции, поэтому используется разработанная авторами технология оценки расстояния, не использующая стереозрение.

в) Для каждого человека, находящегося вблизи рельсового пути, делается оценка расстояния от него до поезда.

г) Используя метод Лукаса-Канаде [10], анализируется движение каждого детектированного человека, чтобы определить, движется ли он в опасную зону, либо успеет выйти из опасной зоны.

В качестве нейронной сети использовалась сеть «mask_rcnn_inception_resnet_v2_1024x1024_coco17» [11, 12], переученная для распознавания рельсового пути и для распознавания людей на фоне железнодорожной инфраструктуры. Сеть функционировала в среде Tensorflow 2 [13].

Программный модуль, анализирующий расстояния и скорости людей вблизи рельсового пути, был написан на C++ с использованием библиотек OpenCV [14].

Для интеграции модуля нейронной сети и программного модуля анализа расстояний и скоростей, в том числе параллельной трансляции видеоряда на оба модуля и организации обмена данными между ними использовалась Robot Operating System версии "Kinetic Kame" [15] под Linux Ubuntu 16.04.

В данной работе реализован новый подход к обучению узкоспециализированной масочной нейронной сети, которая распознает фигуру человека на фоне железнодорожной инфраструктуры и определяет его местоположение относительно рельсовых путей с точки зрения его безопасности. Для создания более качественной маски используются снимки более высокого разрешения, созданные одновременно со снимками штатных камер СТЗ, благодаря чему увеличивается качество распознавания, особенно для больших дистанций, когда объект интереса в сложной фоно-целевой обстановке малозаметен.

Помимо детектирования фигуры человека в окрестности железнодорожных путей нейронная сеть производит быструю оценку расстояния до нее за счет сравнения с близкими предметами, типичные линейные размеры которых заносились в сеть в процессе обучения. Здесь «быстрая оценка» означает, что скорость оценки не менее чем в 3 раза превосходит скорость, обеспечиваемую алгоритмами Block Matching [16] или Semi-Global Block Matching [17]. Соответственно, в отличие от данных алгоритмов, предложенная в данной работе технология не обеспечивает субдециметровую точность, но для решения данной задачи ее точность достаточна.

Для обучения нейронной сети был использован размеченный датасет, состоящий из большого набора (порядка 120 000) фотографий исправных путей без людей-нарушителей, значительно меньшим набором фотографий людей вблизи путей и малого сета фотографий заброшенных путей, находящихся не в лучшем состоянии.

С целью балансировки датасета было сделано следующее: датасет был случайным образом разбит на 6000 примерно равных фрагментов, в каждом фрагменте количество элементов каждого класса было примерно одинаково. Затем с помощью скрипта было инициализировано обучение нейронной сети. Однако для устранения проблемы переобучения обучение выполняется на каждом случайно выбранном фрагменте ровно 10 эпох, после чего завершается и случайным образом выбирается новый фрагмент. На нём также обучается 10 эпох и так далее, пока качество распознавания не будем приемлемым.

Обученная нейронная сеть использована как узел (node) под управлением ROS, то есть она подписана на видеотему от видеокамеры и в ответ на каждое полученное сообщение публикует результаты распознавания как набор битовых масок.

Программный модуль, анализирующий расстояния и скорости людей, написан как узел ROS, подписанный на видеотему от видеокамеры и на результаты распознавания от нейронной сети и публикующий результаты в виде перечисления обнаруженных людей с указанием дистанции и скорости, а также, с целью контроля, в виде изображений.

Программный модуль для анализа расстояний и скорости написан на C++ с использованием библиотек OpenCV, являющихся компонентой ROS «Kinetic Kame» [18].

Оценка исправности рельсового пути делается проверкой, что каждый рельс распознается как связанная область без резких колебаний толщины, а также не имеет резких колебания яркости в продольном направлении.

Рассмотрим подробнее решение задачи оценки расстояния до человека в разработанном алгоритме. На плоском изображении от нижнего края изображения человека (т.е. от его ступней) проводится горизонтальная линия до пересечения с

рельсами включительно и определяется длина в пикселах участка этой линии, оканчивающегося между рельсами (обозначим это число n_1). Затем определим в пикселах длину участка линии от человека до рельса (обозначим это число n_2). Оценкой расстояния в метрах от человека до рельсов будет

$$d_2 = \frac{n_2}{n_1} \cdot d_1, \quad (1)$$

где $d_1 = 1.520$ м, стандарт расстояния между рельсами, принятый в Российских железных дорогах, называемый "ширина колеи" и определяемый как "расстояние между внутренними гранями головок рельсов" [19]. Данное значение сравнивается с результатом «быстрой оценки» расстояния нейронной сетью и корректируется в сторону «быстрой оценки», причем вес коррекции логистически убывает при увеличении разности «быстрой оценки» и d_2 [19].

Кроме этого, мы получим значение

$$L_{pix} = \frac{d_1}{n_1}, \quad (2)$$

указывающее, какое поперечное расстояние в метрах при данном удалении от поезда соответствует одному пикселу.

Оценка расстояния от человека до видеокамеры (то есть, по факту, до поезда), делается 2-мя способами:

а) В ближайшем к человеку участке рельсового пути измеряется видимое горизонтальное расстояние между рельсами в пикселах n_1 , после чего вычисляется приблизительное расстояние от видеокамеры до человека по формуле:

$$d = \frac{A_1}{n_1}, \quad (3)$$

где A_1 полностью определяется оптическими свойствами камеры и шириной колеи и может быть либо вычислена, либо найдена посредством калибровки.

б) Измеряется высота человека в пикселах n_2 и, используя среднюю оценку роста 170 см, приблизительно оценивается расстояние до человека по формуле:

$$d = \frac{A_2}{n_2}, \quad (4)$$

где A_2 полностью определяется оптическими свойствами камеры и может быть либо вычислена, либо найдена посредством калибровки. Берется наименьшее из этих значений.

На рис. 2 на красном фоне оценка расстояния до человека. Оценка не является прецизионной, но точность достаточна для оценки безопасности либо опасности нахождения вблизи путей.

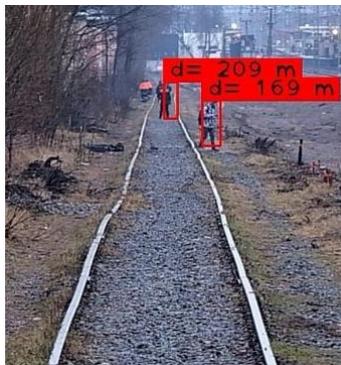


Рис. 2. Примеры определения расстояния до человека возле путей

Оценка скорости человека делается следующим способом. Вначале решим, что "продольная составляющая" скорости, то есть вдоль путей, нам неинтересна, так как она ничего не меняет. Нам интересна именно поперечная составляющая скорости, то есть приближение человека к путям либо его удаление от путей. Поперечную составляющую найдем следующим образом: найдем на изображении видимую поперечную скорость пиксел/сек., после чего, из оценки расстояния до человека, зная соотношение пикселов к метрам на данной дистанции, получим поперечную скорость в метрах в секунду.

Для определения видимой поперечной скорости на изображении применим метод Лукаса-Канаде. А именно, воспользуемся тем, что у нас независимо в разных потоках работают "медленная" нейронная сеть и "быстрая" программа-анализатор расстояния и скорости и к

каждой связной области, выделенной битовой маской как "человек", применим технологию Лукаса-Канаде.

Метод Лукаса-Канаде основан на следующем: пусть у нас есть изображение $I(x, y, t)$, где $I(x, y, t)$ – это яркость как функция пиксельных координат x и y и времени t . Предполагая, что изменение изображения обусловлено его перемещением на малые расстояния δx и δy , происходящим за промежуток времени δt , получим

$$I(x, y, t) = I(x + \delta x, y + \delta y, t + \delta t), \quad (5)$$

где $\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y}, \frac{\partial I}{\partial t}$ – частные производные по координатам x, y и t соответственно; $v_x = \frac{\delta x}{\delta t}, v_y = \frac{\delta y}{\delta t}$ – компоненты скорости оптического потока.

Малые значения величин $\delta x, \delta y$ и δt позволяет нам при разложении функции в ряд Тейлора ограничиться только первыми частными производными по пространственным координатам (x и y) и временной координате t . Это дает возможность свести задачу нахождения оптического потока к решению линейного уравнения относительно двух неизвестных компонент его скорости $\mathbf{v}(v_x, v_y)$:

$$\frac{\partial I}{\partial x} v_x + \frac{\partial I}{\partial y} v_y + \frac{\partial I}{\partial t} = 0, \quad (6)$$

называемого основным уравнением оптического потока.

Очевидно, что уравнение (6) неразрешимо однозначно – возникает так называемая проблема регуляризации, требующая добавления к модели задачи дополнительной информации. Поэтому добавляется требование, чтобы в окрестности "рассматриваемой точки" скорость менялась минимальным образом, что реализуют так: в конечной окрестности "рассматриваемой точки" скорость считают постоянной, в вызванную этим допущение ошибку минимизируют методом наименьших квадратов.

В конечном итоге система уравнений для нахождения компонент скорости приобретает вид:

$$\mathbf{Q}\mathbf{v} = \mathbf{b}, \quad (7)$$

где

$$\mathbf{Q} = \begin{pmatrix} \sum_{i=1}^n q_i \left(\frac{\partial I(b_i)}{\partial x} \right)^2 & \sum_{i=1}^n q_i \frac{\partial I(b_i)}{\partial x} \frac{\partial I(b_i)}{\partial y} \\ \sum_{i=1}^n q_i \frac{\partial I(b_i)}{\partial x} \frac{\partial I(b_i)}{\partial y} & \sum_{i=1}^n q_i \left(\frac{\partial I(b_i)}{\partial y} \right)^2 \end{pmatrix}, \quad (8)$$

$$\mathbf{b} = - \begin{pmatrix} \sum_{i=1}^n q_i \frac{\partial I(b_i)}{\partial x} \frac{\partial I(b_i)}{\partial t} \\ \sum_{i=1}^n q_i \frac{\partial I(b_i)}{\partial y} \frac{\partial I(b_i)}{\partial t} \end{pmatrix}, \quad (9)$$

Здесь b_i – пиксели области, видимая скорость которой нас интересует, q_i – веса пикселей (положительные ненулевые, убывают до нуля на границе области).

В поле скоростей методом наименьших квадратов следует выделить скорость $v_{x1} = v_{x0} + Bx$, $v_{y1} = v_{y0} + By$, обусловленную движением поезда (здесь (v_{x0}, v_{y0}) обусловлена угловой скоростью, (Bx, By) обусловлена движением вперед), то есть (v_{x0}, v_{y0}, B) должны быть выбраны так, чтобы скорость $v - v_1$ большинства фоновых объектов оказались минимальна.

Соответственно, $v - v_1$ и будет пиксельной скоростью. Истинной поперечной скоростью в метрах в секунду будет $(v - v_1)L_{pix}$.

Также в программе оцениваются полученные от нейронной сети данные о местоположении человека и местоположении рельсового пути с целью определения взаимного положения человека и рельсового пути. По результатам оценки маска человека помечается синим цветом, если он в опасной зоне на рельсах и красным цветом, если он находится на краю опасной зоны, но его изображение указывает, что с высокой вероятностью он окажется внутри опасной зоны (человек стоит лицом к пути). Это проиллюстрировано на рис. 3.



Рис. 3. Работа масочной сети

Отметим, что, оценивая не только местоположение человека, но и его контекст (местоположение ближайших объектов) можно оценить поведение попавшего в поле зрения камер СТЗ случайного человека. То есть, кроме низкоуровневой категориальной оценки (например, "человек=90%") нейронная сеть способна дать высокоуровневую категориальную оценку. То есть такая система способна оценить следующую "категориально-когнитивную вероятность":

- ◆ Человек безопасно уйдет до появления поезда.
- ◆ Человек не сможет оценить опасность и ему будет причинен вред в результате столкновения с поездом.
- ◆ В результате действий человека вред будет нанесен инфраструктуре железных дорог либо поезду.

Соответственно, было бы возможным и целесообразным давать временной прогноз ситуации, т.е. либо указывать время ожидаемой реализации опасной ситуации, либо прогноз, что данная ситуация не произойдет по каким-либо причинам. Для этого нейронная сеть должна анализировать траекторию движения объекта по нескольким кадрам. Для этого в используемой нейронной сети должны присутствовать как плоские слои, определяющие общие тенденции и целеполагание ситуаций, так и сверточные слои, в данном случае реализующие в параллельном режиме технологию Лукаса-Канаде построения поля скоростей посредством дифференциального анализа изображений и, тем самым, способные экстраполировать движение объекта и изменение его контекста как функцию времени.

Но, соответственно, кроме высоких требований к архитектуре такой нейронной сети и ее входным данным, не менее высокие требования предъявляются к обучающей выборке данной сети, то есть, чтобы сеть могла делать прогноз и оценивать вероятное время реализации данного прогноза, обучающая выборка должна содержать не статические ситуации, а именно ситуации с ожидаемым прогнозом и известным временем его реализации, причем обучающая выборка должна быть «достаточно полной» а смысле содержания достаточного большого набора ситуаций и прогнозов близких к «условиям эксплуатации сети».

Данная технология будет полезна в областях, где требуется постоянная оценка обстановки и необходимо оперативно решать по местоположению и контексту человека, является ли ситуация штатной, либо потенциально опасной для людей и сооружений и т.д.

Заключение. Системы обнаружения и распознавания движения и объектов повсеместно используются в системах, связанных с повышенной безопасностью движения. Примером таких систем являются автомобильные системы предупреждения о столкновениях и слежения за дорогой. Одной из задач таких систем является обнаружение движущихся объектов, таких как пешеходы, и их движение при переходе проезжей части [20].

Подобные технологии также востребованы и в сфере железнодорожного транспорта, так как данная область связана с четким выполнением заранее построенных графиков движения. Любое отклонение от установленного маршрута может привести к большим экономическим потерям, а несвоевременная оценка обстановки может повлиять на жизни людей. Данная работа содержит основные идеи системы компьютерного зрения, позволяющей обнаруживать препятствия на пути железнодорожного транспорта для их предварительной классификации и принятия решения о дальнейших действиях.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Fioretti F., Ruffaldi E., Avizzano C. A.* A single camera inspection system to detect and localize obstacles on railways based on manifold Kalman filtering // 2018 IEEE 23rd International Conference on Emerging Technologies and Factory Automation (ETFA). – IEEE, 2018. - Vol. 1. – P. 768-775.
2. *Минаков В.А., Фоменко В.К.* Технология машинного зрения на локомотивах для идентификации путевых сигналов // Мир транспорта. – 2020. – Т. 17, №. 6. – С. 62-72.
3. *Mukojima H. et al.* Moving camera background-subtraction for obstacle detection on railway tracks // 2016 IEEE international conference on image processing (ICIP). IEEE, 2016. – P. 3967-3971.
4. *Kyatsandra A.K. et al.* Development of TRINETRA: A Sensor Based Vision Enhancement System for Obstacle Detection on Railway Tracks // IEEE Sensors Journal. – 2022.
5. *He D. et al.* Obstacle detection in dangerous railway track areas by a convolutional neural network // Measurement Science and Technology. – 2021. – Vol. 32, No. 10. – P. 105401.
6. *Sheikh Y., Zhai Y., Shafique K. and Shah M.* Visual monitoring of railroad grade crossing // Proc. SPIE Sensors and Command Control Communications and Intelligence (C3I) Technologies for Homeland Security and Homeland Defense III. – 2004. – Vol. 5403. – P. 654-660.
7. *Konovalenko I.* Overview of methods for estimation the observed velocity of the object in the video stream // Управление, информация и оптимизация (VI ТМШ). – 2014. – С. 34-34.
8. *Le James.* How to do Semantic Segmentation using Deep learning. – URL: <https://nanonets.com/blog/how-to-do-semantic-segmentation-using-deep-learning>.
9. *Lateef F., Ruichek Y.* Survey on semantic segmentation using deep learning techniques // Neurocomputing. – 2019. – Vol. 338. – P. 321-348.
10. *Визильтер Ю.В., Желтов С.Ю., Бондаренко А.В. и др.* Обработка и анализ изображений в задачах машинного зрения. – М.: Физматкнига, 2010. – 672 с.
11. *He K. et al.* Mask r-cnn // Proceedings of the IEEE international conference on computer vision. – 2017. – P. 2961-2969.

12. TensorFlow 2 Detection Model Zoo. – URL: https://github.com/tensorflow/models/blob/master/research/object_detection/g3doc/tf2_detection_zoo.md.
13. Gulli A., Kapoor A., Pal S. Deep learning with TensorFlow 2 and Keras: regression, ConvNets, GANs, RNNs, NLP, and more with TensorFlow 2 and the Keras API. – Packt Publishing Ltd, 2019.
14. Bradski G., Kaehler A. OpenCV // Dr. Dobb's journal of software tools. – 2000. – Т. 3. – С. 2.
15. ROS Kinetic Kame. - URL: <http://wiki.ros.org/kinetic>.
16. Chang Q., Maruyama T. Real-time stereo vision system: a multi-block matching on GPU // IEEE Access. – 2018. – Vol. 6. – P. 42030-42046.
17. Banz C., Blume H., Pirsch P. Real-time semi-global matching disparity estimation on the GPU // 2011 IEEE International Conference on Computer Vision Workshops (ICCV Workshops). – IEEE, 2011. – P. 514-521.
18. Fairchild C., Harman T.L. ROS Robotics By Example: Learning to control wheeled, limbed, and flying robots using ROS Kinetic Kame. – Packt Publishing Ltd, 2017.
19. Карпачевский В.В. Правила технической эксплуатации железных дорог. – 2017.
20. Sam Schauland, Joerg Velten, Anton Kummert. Motion-Based Object Detection for Automotive Applications using Multidimensional Wave Digital Filters // VTC Spring 2008 - IEEE Vehicular Technology Conference, Singapore, Singapore, 20 May 2008. – P. 2700-2704.

REFERENCES

1. Fioretti F., Ruffaldi E., Avizzano C. A. A single camera inspection system to detect and localize obstacles on railways based on manifold Kalman filtering, 2018 IEEE 23rd International Conference on Emerging Technologies and Factory Automation (ETFA). IEEE, 2018, Vol. 1, pp. 768-775.
2. Minakov V.A., Fomenko V.K. Tekhnologiya mashinnogo zreniya na lokomotivakh dlya identifikatsii putevykh signalov [Technology of machine vision on locomotives for identification of track signals], *Mir transporta* [Mir transport.], 2020, Vol. 17, No. 6, pp. 62-72.
3. Mukojima H. et al. Moving camera background-subtraction for obstacle detection on railway tracks, 2016 IEEE international conference on image processing (ICIP). IEEE, 2016, p. 3967-3971.
4. Kyatsandra A.K. et al. Development of TRINETRA: A Sensor Based Vision Enhancement System for Obstacle Detection on Railway Tracks, *IEEE Sensors Journal*, 2022.
5. He D. et al. Obstacle detection in dangerous railway track areas by a convolutional neural network, *Measurement Science and Technology*, 2021, Vol. 32, No. 10, pp. 105401.
6. Sheikh Y., Zhai Y., Shafique K. and Shah M. Visual monitoring of railroad grade crossing, *Proc. SPIE Sensors and Command Control Communications and Intelligence (C3I) Technologies for Homeland Security and Homeland Defense III*, 2004, Vol. 5403, pp. 654-660.
7. Konovalenko I. Overview of methods for estimation the observed velocity of the object in the video stream, *Upravlenie, informatsiya i optimizatsiya (VI TMSH)* [Management, Information and optimization (VI TMSH)], 2014, pp. 34-34.
8. Le James. How to do Semantic Segmentation using Deep learning. Available at: <https://nanonets.com/blog/how-to-do-semantic-segmentation-using-deep-learning>.
9. Lateef F., Ruichek Y. Survey on semantic segmentation using deep learning techniques, *Neurocomputing*, 2019, Vol. 338, pp. 321-348.
10. Vizil'ter Yu.V., Zheltov S.Yu., Bondarenko A.V. i dr. Obrabotka i analiz izobrazheniy v zadachakh mashinnogo zreniya [Image processing and analysis in machine vision problems]. Moscow: Fizmatkniga, 2010, 672 p.
11. He K. et al. Mask r-cnn, *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2961-2969.
12. TensorFlow 2 Detection Model Zoo. Available at: https://github.com/tensorflow/models/blob/master/research/object_detection/g3doc/tf2_detection_zoo.md.
13. Gulli A., Kapoor A., Pal S. Deep learning with TensorFlow 2 and Keras: regression, ConvNets, GANs, RNNs, NLP, and more with TensorFlow 2 and the Keras API. Packt Publishing Ltd, 2019.
14. Bradski G., Kaehler A. OpenCV, *Dr. Dobb's journal of software tools*, 2000, Vol. 3, pp. 2.
15. ROS Kinetic Kame. - URL: <http://wiki.ros.org/kinetic>.

16. Chang Q., Maruyama T. Real-time stereo vision system: a multi-block matching on GPU, *IEEE Access*, 2018, Vol. 6, pp. 42030-42046.
17. Banz C., Blume H., Pirsch P. Real-time semi-global matching disparity estimation on the GPU, *2011 IEEE International Conference on Computer Vision Workshops (ICCV Workshops)*. IEEE, 2011, pp. 514-521.
18. Fairchild C., Harman T.L. ROS Robotics By Example: Learning to control wheeled, limbed, and flying robots using ROS Kinetic Kame. Packt Publishing Ltd, 2017.
19. Karpachevskiy V.V. Pravila tekhnicheskoy ekspluatatsii zheleznikh dorog [Rules of technical operation of railways], 2017.
20. Sam Schauland, Joerg Velten, Anton Kummert. Motion-Based Object Detection for Automotive Applications using Multidimensional Wave Digital Filters, *VTC Spring 2008 - IEEE Vehicular Technology Conference, Singapore, Singapore, 20 May 2008*, pp. 2700-2704.

Статью рекомендовал к опубликованию к.ф.-м.н. С.В. Семендяев.

Шишков Дмитрий Леонидович – МФТИ; e-mail: shishkov_dmitry_geor@mail.ru; г. Долгопрудный, Россия; ведущий программист-разработчик лаборатории волновых процессов и систем управления.

Зарипов Михаил Нилович – e-mail: mikezmn@mail.ru; ведущий программист-разработчик лаборатории волновых процессов и систем управления.

Горбачев Роман Александрович – к.т.н.; зав. лабораторией волновых процессов и систем управления.

Shishkov Dmitry Leonidovich – MIPT; e-mail: shishkov_dmitry_geor@mail.ru; Dolgoprudnyy, Russia; leading programmer and developer of the laboratory of wave processes and control systems.

Zaripov Mikhail Nilovich – e-mail: mikezmn@mail.ru; leading programmer and developer of the laboratory of wave processes and control systems.

Gorbachev Roman Alexandrovich – cand. of eng. sc.; head of the laboratory of wave processes and control systems

УДК 007:621.865.8

DOI 10.18522/2311-3103-2022-1-268-278

В.П. Носков, А.Н. Курьянов

ИСПОЛЬЗОВАНИЕ КОМПЛЕКСИРОВАННЫХ ДЕСКРИПТОРОВ В РЕШЕНИИ SLAM-ЗАДАЧИ

Рассмотрена актуальная задача определения всех шести координат (трех линейных и трех угловых) текущего положения мобильного робота (беспилотного летательного аппарата) по видео-дальномерным изображениям внешней среды (объемным раскрашенным облакам точек), формируемым бортовой комплексированной системой технического зрения, построенной на базе 3D-дальномерного сенсора (лидара) и цветной видеокамеры, при движении (полете) в неизвестной среде. Предложен алгоритм видеонавигации, основанный на использовании комплексированных (видео-дальномерных) дескрипторов, для описания которых используются яркостные и геометрические параметры. Сформулированы правила формирования комплексированного дескриптора, обеспечивающие выделение с помощью оператора Собеля особых (центральных) точек дескриптора и вычисление яркостных и геометрических параметров в его локальной области. Дополнение яркостных параметров дескриптора, формируемых видеокамерой, геометрическими параметрами, формируемых дальномерным сенсором, снимает проблему инвариантности дескриптора к масштабу и тем самым существенно снижает трудоемкость вычислений при его выделении. Описаны правила нахождения соответствующих друг другу комплексированных дескрипторов в последовательности комплексированных изображений, основанные на вычислении разности яркостных и геометрических парамет-