Оригинальная статья
https://doi.org/10.21683/1729-2646-2024-24-3-24-33

Применение преобразования Хафа для определения границы путей в задачах компьютерного зрения аппаратно-программного комплекса фиксации исполненного движения

Using Hough transform to define track boundaries as part of machine vision by train traffic-tracking hardware and software systems

Полевский И.С.
Polevskiy I.S.
Научно-исследовательский и проектно-конструкторский институт информатизации, автоматизации и связи на железнодорожном транспорте (АО «НИИАС»), Российская Федерация, Москва
Research and Design Institute for Information Technology, Signalling, and Telecommunications in Railway Transportation (JSC NIIAS), Russian Federation, Moscow
i.polevskiy@vniias.ru

Полевский И.С.

Резюме. Цель. *При решении задач компьютерного зрения для определения границ детектируемого объекта, как правило, используются методы семантической сегментации, которые требуют высокого вычислительного ресурса. Их использование повышает сложность реализации и увеличивает стоимость решений для внедряемых аппаратно-программных комплексов. В настоящей работе предлагается альтернативный метод определения границы сегментируемого объекта, в виде железнодорожного пути, для комплекса фиксации исполненного движения.* **Методы.** *Так как железнодорожный путь на изображении можно представить линией полинома n-порядка, то для решения задачи детектирования границы пути предлагается использовать приближения в виде прямых линий. В качестве метода детектирования прямых линий предлагается использовать преобразование Хафа, параметрическое пространство которого будет скомпоновано в соответствии с решаемой задачей.* **Заключение.** *Предложенная аппроксимация позволит отказаться от семантической сегментации и снизит вычислительную сложность нагрузки на аппаратуру.*

Abstract. Aim. *In the context of machine vision, the problem of detected object boundary definition is usually solved using semantic segmentation that requires a high computational resource. Its application increases the complexity and the cost of the implemented hardware and software systems. This paper proposes an alternative method for defining the boundary of a segmented object, i.e., a railway track, for a train traffic tracking system.* **Methods.** *Since a railway track in an image can be represented by an n-th order polynomial, it is suggested to solve the problem of track boundary detection by using approximations in the form of straight lines. It is suggested using the Hough transform to detect straight lines. The former's parametric space will be arranged in accordance with the problem to be solved.* **Conclusions.** *The proposed approximation will allow abandoning semantic segmentation and reduce the computational complexity and load.*

Ключевые слова: *фиксация исполненного движения, компьютерное зрение, машинное обучение, преобразование Хафа, семантическая сегментация.*
Keywords: *train traffic tracking, machine vision, machine learning, Hough transform, semantic segmentation.*

Для цитирования: *Полевский И.С. Применение преобразования Хафа для определения границы путей в задачах компьютерного зрения аппаратно-программного комплекса фиксации исполненного движения // Надежность. 2024. №3. С. 24-33. https://doi.org/10.21683/1729-2646-2024-24-3-24-33*
For citation: *Polevskiy I.S. Using Hough transform to define track boundaries as part of machine vision by train traffic-tracking hardware and software systems. Dependability 2024;3:24-33. https://doi.org/10.21683/1729-2646-2024-24-3-24-33*

Поступила: *03.11.2023* / **После доработки:** *14.05.2024* / **К печати:** *09.09.2024*
Received on: *03.11.2023* / **Revised on:** *14.05.2024* / **For printing:** *09.09.2024*

Введение

АО «НИИАС» разрабатывает Аппаратно-программный комплекс фиксации исполненного движения, далее – АПК ФИД. Технологической задачей разрабатываемого комплекса является автоматическая фиксация приема, отправления и пропуска поездов на станции с помощью средств искусственного интеллекта: решение задач классификации объектов и задач регрессии. Главной целью внедрения комплекса АПК ФИД является реализация и использование методов, которые позволят отказаться от средств и устройств электрической централизации на станции, и оперировать только данными, которые поступают автоматически от устройств оптического видеоконтроля.

Одной из технических обеспечивающих задач комплекса АПК ФИД является задача определения границы железнодорожного пути на изображениях анализируемого видеоряда. Решение такого рода задач, как правило, осуществляется с помощью методов семантической сегментации. Семантическая сегментация – это решение, при котором обеспечивается детальная классификация объекта, где происходит присвоение семантических меток классов для каждого пикселя изображения.

Современные методы семантической сегментации основаны на глубоких сверточных сетях [1]. Например, метод семантической сегментации PSPNet [2] реализован на базе сверточной сети ResNet [3], а метод DeepLabv3 [4] основан на модели сети xception [5]. Применение сверточных сетей в методах семантической сегментации показывает высокую точность сегментирования объектов [2, 4, 6], но имеет и высокую вычислительную сложность, которая напрямую зависит от выбранной архитектуры сверточной сети, количества сверточных слоев и числа обучаемых параметров. При сравнительном анализе [7] применения популярных методов семантической сегментации изображений (PSPNet, DeepLabv3, UNet [8]) на внедряемых аппаратных платформах NVIDIA Jetson было показано, что генерация семантической маски для одного изображения с разрешением 4000 × 3000 точек может занимать до нескольких секунд. Это ограничивает применение подобных методов для внедряемого комплекса АПК ФИД в силу ограниченной вычислительной мощности оборудования.

Целью данной работы является введение альтернативного метода определения границы железнодорожного пути на изображении. Предлагается использовать метод, который построен на вычислительном алгоритме преобразования Хафа [9]. Данный метод основан на семантике границы детектируемого объекта с помощью геометрических примитивов в виде прямых линий.

Рассмотрение метода

Железнодорожный путь ограничен рельсами, каждый из которых можно представить в виде линии полинома n-порядка при линейной перспективе обрабатываемого изображения. Степень такого полинома зависит от кривизны пути на заданном масштабе изображения в кадре, где масштаб определяется высотой установки видеокамеры, ее угла обзора и фокусного расстояния. Таким образом, с помощью преобразования Хафа можно выделить границу железнодорожного пути набором прямых линий, определяющих ломаную кривую. Такая ломаная кривая будет являться дискретным представлением исходного полинома линии, который можно получить путем интерполяции с помощью численных методов решения задачи приближения функций[1]. На рис. 1 показан пример определения границы железнодорожного пути с помощью преобразования Хафа.

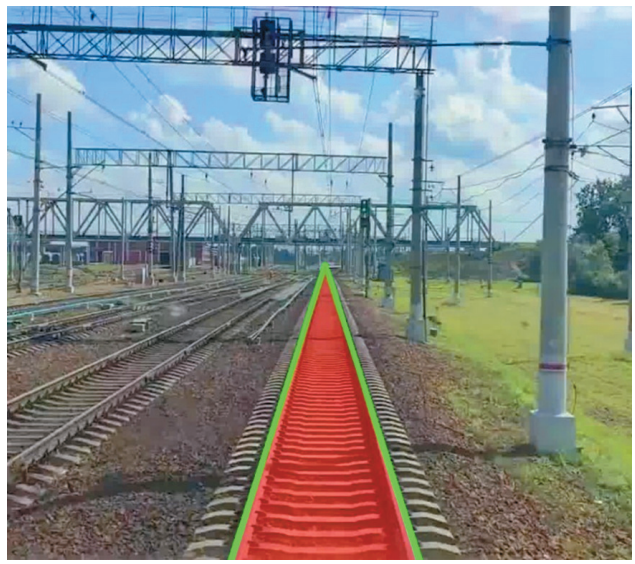

Рис. 1. Пример определения границы железнодорожного пути с помощью преобразования Хафа (красным цветом выделен путь, зеленым цветом показаны прямые линии, определяющие его границы)

Алгоритм преобразования Хафа был запатентован Полем Хафом в 1962 году [9]. Преобразование Хафа применяется для поиска геометрических фигур на растровых изображениях. Алгоритм основан на параметризации пикселей с помощью полярных координат: угла наклона и длины радиус-вектора от заданного центра анализируемого изображения. Результатом преобразования является Хаф-образ (параметрическое пространство). Построение Хаф-образа происходит по формуле (1)

$$\rho = x\cos\theta + y\sin\theta, \quad (1)$$

где ρ – длина радиус-вектора, θ – угол наклона, x – координата пикселя по оси абсцисс, y – координата пикселя по оси ординат.

Класс детектируемой фигуры определяется путем голосования в параметрическом пространстве, где суммируются значения для пикселей с одинаковыми по-

[1] Даутов Р.З., Тимербаев М.Р. Численные методы. Приближение функций: учебное пособие. Казань: Казан. ун-т, 2021. 123 с.

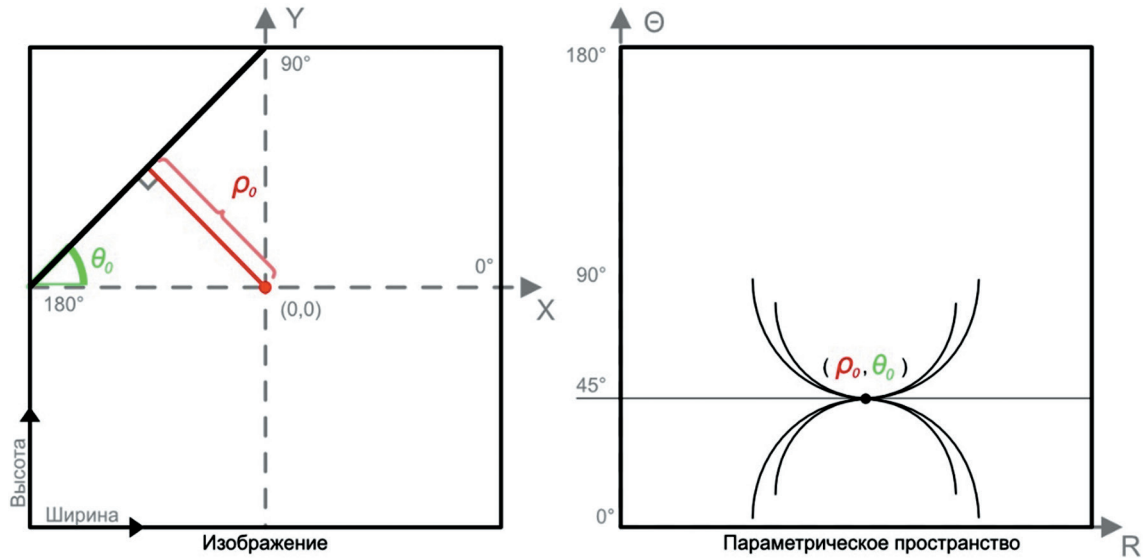

Рис. 2. Визуализация преобразования Хафа для прямой линии

лярными координатами. Голосование выделяет локальные максимумы Хаф-образа, которые задают паттерн искомого объекта. Любая прямая линия на исходном графическом изображении представляется точкой локального максимума в параметрическом пространстве. На рис. 2 показана визуализация преобразования Хафа для прямой линии.

Компоновка пространства Хафа

Преобразование Хафа имеет высокую асимптотическую сложность – $O(N^3)$, а само параметрическое пространство занимает большой объем памяти, так как для каждого пикселя изображения строится вектор значений по каждому рассматриваемому углу наклона θ на промежутке $[0; \pi)$.

На практике существуют методы, позволяющие ускорить построение параметрического пространства Хафа или его приближения. В 1992 году впервые был предложен метод быстрого преобразования Хафа [10], далее – БПХ, который снижает вычислительную сложность построения Хаф-образа за счет использования факта самопересечения рассматриваемых дискретных прямых. Асимптотическая сложность алгоритма БПХ составляет $O(N^2 \log N)$. Также широкое применение находят алгоритмы предварительной обработки растровых изображений, целью использования которых является снижение количества параметризуемых пикселей (бинаризация изображения): оператор Кэнни [11], оператор Собеля [12], алгоритм Брэдли-Рота [13] и др.

Предполагается, что предлагаемый метод будет применяться в составе вычислительного графа предиктивной модели комплекса АПК ФИД, где на вход алгоритма будут переданы бинаризированные карты признаков. Математическим ядром метода выбран классический алгоритм преобразования Хафа, который, в сравнении с БПХ, обладает рядом ключевых преимуществ для решаемой задачи:

1) не имеет рекурсивных операций. Использование рекурсивных операций в общем случае не желательно при параллельных вычислениях на графическом процессоре, требуются их дополнительные преобразования к итеративному виду [14], что может увеличить расход памяти и существенно усложнить конечную реализацию алгоритма;

2) имеет высокую точность работы. В алгоритме БПХ прямая линия параметризуется диадическим шаблоном, который является ее дискретным приближением в рамках растрового изображения, а для классического преобразования Хафа прямая линия – это один параметризованный пиксель. Следовательно, в определенных случаях классический алгоритм имеет более высокую точность работы, так как в нем не используются приближенные методы вычислений;

3) обладает простотой реализации на вычислительном графе. Обучение модели с использованием алгоритма БПХ будет затрудненно в виду специфики построения его параметрического пространства: необходимо сохранять всю историю формирования диадических шаблонов согласно условиям и циклам рекурсивного алгоритма.

В рамках данной работы предлагается привести параметрическое пространство Хафа к компактному дискретному виду в заданном диапазоне значений – выполнить его квантование.

Дискретное представление параметрического пространства Хафа предполагает использование интервалов квантования для радиус-вектора ($\Delta\rho$) и для угла наклона ($\Delta\theta$) – формула (2)

$$\hat{\rho} = \frac{\rho}{\Delta\rho}; \hat{\theta} = \frac{\theta}{\Delta\theta}, \quad (2)$$

где $\hat{\rho}$ – квантованный параметр радиус-вектора, $\hat{\theta}$ – квантованный параметр угла наклона, ρ – длина радиус-вектора, θ – угол наклона, $\Delta\rho$ – заданный интервал квантования радиус-вектора, $\Delta\theta$ – заданный интервал квантования угла наклона.

Выбор интервалов квантования напрямую влияет на время построения Хаф-образа и определяет точность детектирования прямых линий, которые будут описывать границы железнодорожного пути на исходном изображении. В комплексе АПК ФИД предполагается установка видеокамер на ригели, которые проходят над путями. Следовательно, получаемое изображение с камеры будет иметь точку линейной перспективы, а прямые линии – неоднородные искажения, зависящие от ориентации и расстояния к центру перспективы. Каждая ведеокамера имеет заданную зону интереса, в рамках которой осуществляется контроль расположенных там путей. Таким образом, для решаемой задачи можно выделить два варианта интервалов квантования: линейное и нелинейное. Для линейного квантования будет выбран фиксированный шаг, а для нелинейного подбор шага будет осуществляться в соответствии с зоной интереса и контроля видеокамеры.

Так как длина радиус-вектора ρ определяет смещение прямой линии относительно заданного центра координат изображения, то ошибка интерполяции параметра $\hat{\rho}$ не приведет к пересечению искомой прямой линии и фактической границы пути. Следовательно, для параметра ρ выбран линейный интервал квантования. В таком случае количество уровней квантования (размерность пространства) будет вычисляться по формуле (3)

$$R = \frac{\sqrt{H^2 + W^2}}{\Delta\rho}, \tag{3}$$

где R – количество уровней квантования, H – высота изображения, W – ширина изображения, Δρ – заданный интервал квантования радиус-вектора.

Для угла θ вводятся линейный и нелинейный виды интервалов квантования. Для линейного интервала квантования угла θ количество уровней квантования будет вычисляться в соответсвии с формулой (4)

$$\Theta = \frac{\pi}{\Delta\theta}, \tag{4}$$

где Θ – количество уровней квантования, Δθ – заданный интервал квантования угла.

Нелинейный интервал квантования угла θ предполагает выделение зоны интереса, для которой будет повышена разрешающая способность параметрического пространства Хафа. В таком случае число уровней квантования будет задано в соответствии с формулой (4), а характер распределения будет задан логарифмическим законом.

В связи с тем, что в комплексе АПК ФИД анализируемое графическое изображение обладает линейной перспективой, а железнодорожные пути, находящиеся в зоне интереса, расположены близко к вертикальной оси изображения, повышение разрешающей способности будет осуществляется логарифмически от 0 до 90 градусов. Таким образом, предлагается увеличить точность детектирования близких к вертикальным прямым линиям описывающих границу железнодорожного пути на изображении. Пример определения зоны интереса показан на рис. 3.

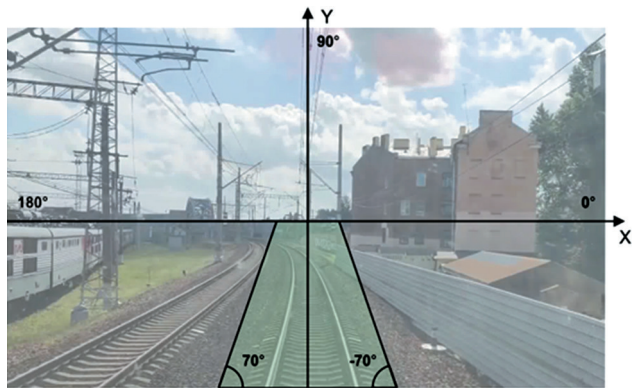

Рис. 3. Пример определения зоны интереса (зона интереса выделена зеленым цветом и определена для прямых линий изображения, угол наклона которых находится в интервале от 90 до 70 градусов)

Для логарифмического интервала квантования вводится основание логарифма, определяющее характер распределения при условии $\theta \in [0; \frac{\pi}{2}]$ – формула (5)

$$b = (\frac{\gamma\pi}{2} + 1)^{\frac{2}{\Theta}}, \tag{5}$$

где b – расчетное основание логарифма, γ – коэффициент масштабирования, Θ – количество уровней квантования угла.

Функция расчета квантованного значения параметра $\hat{\theta}$ определяется формулой (6), а обратная ей функция интерполяции для расчета угла наклона θ определяется в соответствии с формулой (7)

$$f(\theta) = \frac{\Theta}{2} - \log_b(\gamma(\frac{\pi}{2} - \theta) + 1), \tag{6}$$

$$f^{-1}(\hat{\theta}) = \frac{\pi}{2} - \frac{b^{(\frac{\Theta}{2} - \hat{\theta})} - 1}{\gamma}, \tag{7}$$

где θ – угол наклона прямой, $\hat{\theta}$ – квантованное значение угла наклона прямой, Θ – количество уровней квантования угла, b – расчетное основание логарифма, γ – коэффициент масштабирования.

Следует отдельно выделить, что зона интереса применяется для всего анализируемого изображения, то есть логарифмическое распределение будет применено для всех углов прямых линий изображения. Соответственно, низкая разрешающая способность для линий, близких к горизонтальным, даст высокую ошибку интерполяции по сравнению с линейным интервалом квантования углов.

Сравнительный анализ

Конфигурация дискретного параметрического пространства Хафа определяется его гиперпараметрами: размеренностью пространства Θ и R, а также коэф-

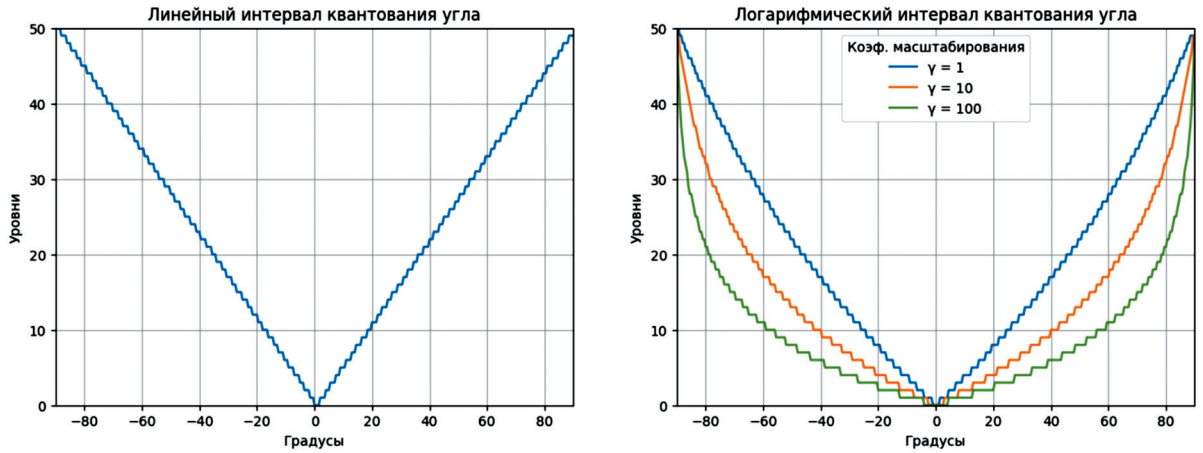

Рис. 4. Разница в распределениях параметризованного угла наклона θ при линейном и нелинейном интервалах квантования (для шага угла 0,5 и заданных гиперпараметров: Θ=100, γ∈{1,10,100})

фициентом масштабирования γ в случае нелинейного интервала квантования. Разница в распределениях параметризованного угла наклона $\hat{\theta}$ при линейном и нелинейном интервалах квантования показана на рис. 4.

Для количественной оценки ошибки интерполяции необходимо выбрать метод, с помощью которого можно сравнить прямые линии.

В задачах компьютерного зрения широкое применение находят метрики CD (Chamfer distance) [15] и IoU (Intersection Over Union) [16]. Данные метрики работают на уровне пикселей обрабатываемого изображения. Метрика CD оценивает сходство между двумя наборами точек, а метрика IoU дает оценку отношениям площадей детектируемых объектов.

Существует специализированная версия метрики IoU – mIoU [17], которая адаптирована для сравнения прямых линий на изображении. Такая метрика оценивает площадь пересечения двух сравниваемых прямых линий к общей площади изображения согласно формуле 8.

$$mIoU(l_1, l_2, I) = \frac{S(l_1 \bigcap l_2)}{S(I)}, \quad (8)$$

где $mIoU$ – оценочная метрика, l_1 – первая прямая линия, l_2 – вторая прямая линия, I – изображение, S – функция расчета площади.

Особенностью метрики mIoU является ее диспропорция: для двух пересекающихся прямых линий, между которыми задан фиксированных угол наклона, метрика mIoU будет показывать разное значение, зависящее от конкретного места расположения прямых на изображении. Данная метрика удобна при относительном сравнении линий на растровых изображениях.

На рис. 5 показано сравнение двух прямых линий с помощью метрики mIoU. Угол между прямыми линиями фиксирован и составляет 10 градусов. Видно, что при изменении ориентирования прямых на изображении метрика mIoU имеет разное значение, напрямую зависящее от площади пересечения.

При обработке цифровых сигналов в контексте задач квантования, используется термин «шум квантования»

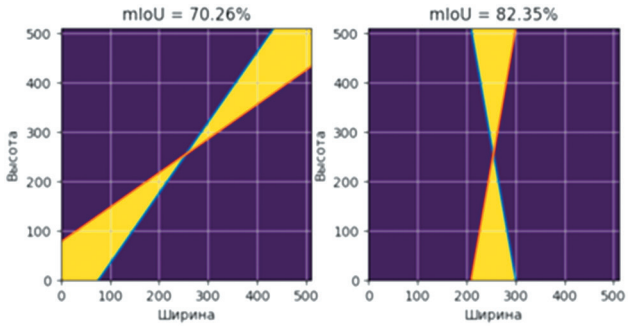

Рис. 5. Сравнение двух прямых линий с помощью метрики mIoU (площадь пересечения прямых показана желтым цветом)

[18], который представляет собой величину ошибки, возникающую в сигнале в результате процесса квантования. Эту ошибку можно использовать для оценки точности интерполяции, как абсолютную – без учета разрешения изображения, количества его пикселей на единицу площади. Расчет ошибки квантования производится по формуле (9)

$$e_i = \hat{x}_i - x_i, \quad (9)$$

где e_i – ошибка квантования, $\hat{x}$ – значение квантованного отсчета сигнала, x – истинное значение сигнала, i – номер отсчета.

Так как изменение параметра ρ влияет на смещение прямой линии относительно центра изображения, а интерполяция параметра $\hat{\rho}$ не приводит к пересечению искомой линии и детектируемой границы пути, то при оценке сходства прямых линий для данного параметра задано фиксированное значение, соответствующее центру изображения. А оценка точности интерполяции производится только для угла наклона линий θ по отношению к 1 градусу и представлена в процентах согласно формуле (10)

$$E_i = \frac{180 * e_i}{\pi} \cdot 100, \quad (10)$$

где E_i – нормированная ошибка квантования в процентах, e_i – ошибка квантования, i – номер отсчета.

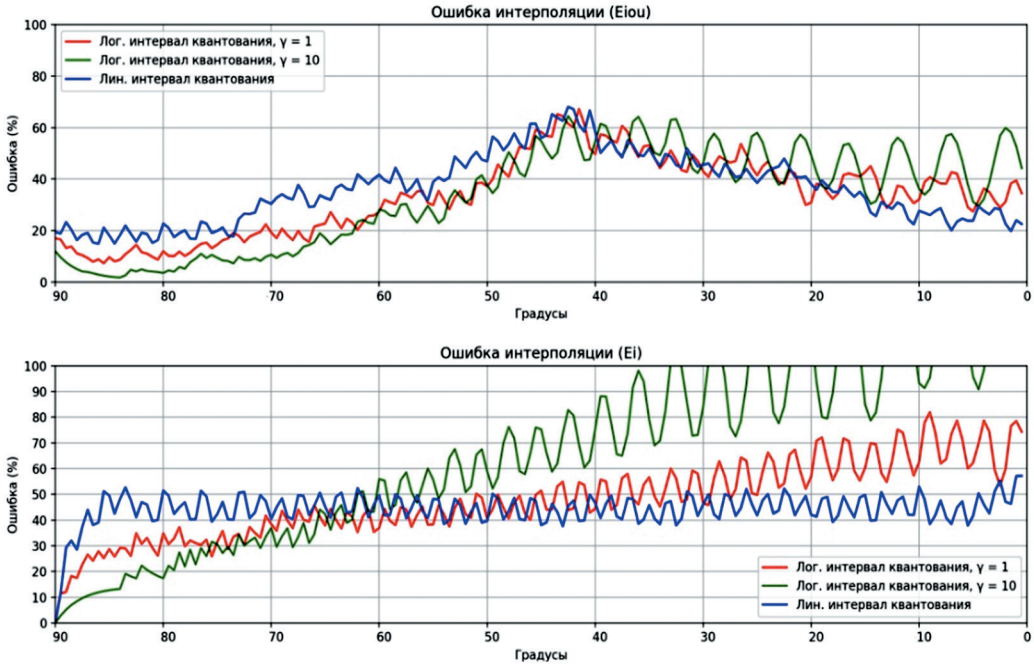

Рис. 6. Графики распределения ошибок интерполяции E_{iou} и E_i параметра $\hat{\theta}$ на интервале от 0 до 90 градусов (размеренность исходного графического изображения 512 × 512 точек, шага угла 0,5, Θ=100, R=100, $\gamma \in \{1,10\}$)

Для сравнительного анализа точности интерполяции, значения ее ошибки, проведена адаптация метрики mIoU – нормирование, в соответствии с формулой (11)

$$E_{iou} = (1 - \frac{mIoU - min(mIoU)}{\max(mIoU) - \min(mIoU)}) \cdot 100, \qquad (11)$$

где E_{iou} – нормированная ошибка $mIoU$ в процентах, $mIoU$ – оценочная метрика.

На рис. 6 приведены графики распределения ошибок интерполяции E_{iou} и E_i параметра $\hat{\theta}$ на интервале от 0 до 90 градусов. Следует отметить, что подбор гиперпараметров Θ и R параметрического пространства Хафа для анализа был произведен таким образом, чтобы размах абсолютного значения метрики mIoU был в рамках одного процента для обоих рассматриваемых интервалов квантования угла θ. Выбранная размеренность исходного графического изображения для метрики mIoU составила 512 × 512 точек.

Таким образом, при рассмотрении графиков распределения ошибок интерполяции для линейного и нелинейного интервалов квантования видно, что нелинейный интервал квантования позволяет получить более высокую точность определения прямых линий для заданной зоны интереса. В табл. 1 показана разница между ошибками интерполяции для рассматриваемых интервалов квантования.

Из представленной табл. 1 видно, что конфигурация зоны интереса может быть осуществлена с помощью

Табл. 1. Разница между ошибками интерполяции для рассматриваемых интервалов квантования

Угол (град.)	Линейный интервал квантования	Нелинейный интервал квантования			
	E_i, %	γ=1		γ=10	
		E_i, %	ΔE_i	E_i, %	ΔE_i
0	45,76	59,43	–29,87%	105,32	–130,16%
10	53,04	62,53	–17,89%	93,02	–75,72 %
20	42,08	56,88	–35,17%	113,93	–170,75%
30	46,12	46,05	–0,82%	84,00	–82,13%
40	46,71	44,37	**5,01%**	75,28	–61,16%
50	39,68	40,66	–2,47%	59,06	–48,84%
60	49,63	36,95	**25,55%**	55,94	–12,71%
65	45,23	37,83	**16,36%**	42,36	**6,35%**
70	40,44	38,47	**4,87%**	36,95	**8,63%**
75	42,25	31,45	**25,56%**	30,63	**27,5%**
80	51,05	34,84	**32,35%**	17,39	**66,23%**
85	49,54	28,71	**42,05%**	12,76	**74,24%**
89	29,28	12,06	**58,81%**	5,17	**82,34%**

параметра γ, который определяет коэффициент масштабирования для нелинейного интервала квантования. Увеличение коэффициента γ приводит к повышению разрешающей способности параметрического пространства для детектируемых прямых линий с углом наклона, близким к 90 градусам, и к уменьшению разрешающей способности при углах, близких к 180 градусам. Разница в ошибках интерполяции ΔE_i показана в процентном отношении между линейным и нелинейным интервалами квантования, где жирным шрифтом выделен прирост разрешающей способности для заданного коэффициента масштабирования γ.

Эксперимент на реальных данных

Для проведения эксперимента на реальных данных выбрана аппаратная платформа NVIDIA Jetson Nano исполнения B01 с общей вычислительной мощностью 0,5 терафлопс. Обучение предиктивных моделей производилось на видеокарте NVIDIA GeForce 1050Ti. В качестве обучающего набора данных использована видеозапись с регистратора смонтированного в кабине машиниста поезда. Для аннотации данных разработано специализированное программное обеспечение (рис. 7). Все изображения обучающего набора данных приведены к разрешению 1280 × 720 точек, устранены радиальная и тангенциальная дисторсии. Таким образом, для обучения предиктивных моделей подготовлено 10 000 изображений со следующим распределением: зима – 48%, лето – 52%. В рамках эксперимента выбраны изображения с прямым профилем пути.

В качестве классического метода семантической сегментации изображений для сравнительного анализа выбрана сверточная сеть UNet [8]. Входом модели сети является изображение размером 320 × 320 точек, а требуемый результат работы предиктивной модели – бинарная маска, где каждый пиксель изображения обозначен как принадлежащий или не принадлежащий к железнодорожному пути. Следовательно, модель сети UNet адаптирована к бинарной классификации и на выходе модели использована логистическая функция согласно формуле (12)

$$\delta(z) = \frac{1}{(1+e^{-z})}, \qquad (12)$$

где δ – вероятностная оценка, z – исходное значение каждого пикселя карты свойств.

Для предлагаемого в рамках работы метода разработана предиктивная модель сверточной сети с условным наименованием HoughNet. Входом модели сети является изображение с размерностью 320 × 320 точек. Модель HoughNet использует сверточные слои энкодера сети UNet для создания бинаризированных карт признаков исходного изображения. Для достижения инвариантности при определении искомых прямых линий, карты признаков берутся в трех масштабах: 160 × 160, 80 × 80, 40 × 40. Далее каждая карта со своей размерностью проходит через соответствующий слой сети «Hough», на котором реализовывается преобразование Хафа и билинейная интерполяция карты к заданному дискретному параметрическому пространству Хафа: Θ=100 и R=100. Результатом работы модели является конкатенация карт признаков, пропущенная через логистическую функцию (формула (12)) – предсказанный Хаф-образ.

Критерием ошибки при обучении обоих предиктивных моделей выбран результат работы функции потерь бинарной кросс-энтропии – формула (13)

$$loss = -\frac{1}{N}\sum_{i=1}^{N}(y_i \log_2 P(y_i) + (1-y_i)\log_2(1-P(y_i))), \qquad (13)$$

где $loss$ – ошибка, y_i – принадлежность к одному из двух классов, $P(y_i)$ – вероятность для первого класса, $(1–P(y_i))$ – вероятность для второго класса, N – число пикселей на изображении, $1/N$ – показатель равновероятного распределения классов для каждого пикселя изображения.

На рис. 8 показана структурная схема используемых предиктивных моделей.

Подготовлено 750 изображений для имитации видеоряда продолжительностью 0,5 минуты при 25 кадрах в секунду, которые содержат три железнодорожных пути с прямым профилем, где каждый путь размечен человеком для последующего сравнительного анализа и оценок. Для оценки работы предиктивных моделей выбрана метрика mIoU. Так как при семантической сегментации возможны разрывы в предсказанной бинарной маске, что связано с объектами на путях, то площадь такого разрыва будет просуммирована с площадью пересечения в рамках оценочной метрики mIoU для обоих прямых линий, характеризующих границу железнодорожного пути. По результатам проведения эксперимента составлена табл. 2, в рамках которой сведены данные по работе обеих рассматриваемых предиктивных моделей на аппаратной платформе NVIDIA Jetson Nano.

Табл. 2 отражает среднюю метрику mIoU для каждого из трех сегментированных путей на изображении, которые отсортированы в следующем порядке следования: «1-ый путь» – линии границы железнодорожного пути, чей угол наклона наиболее близок к 90 градусам, тогда

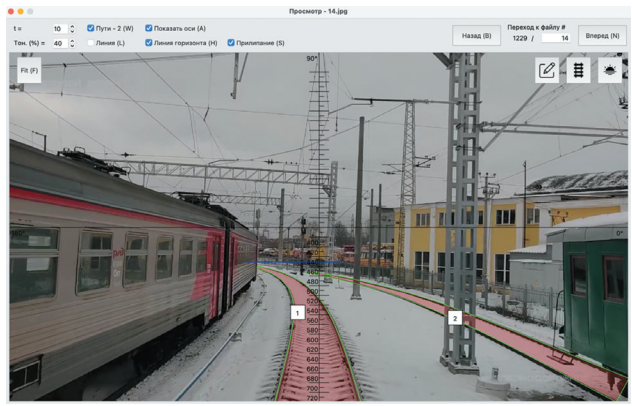

Рис. 7. Программное обеспечение для аннотации обучающего набора данных

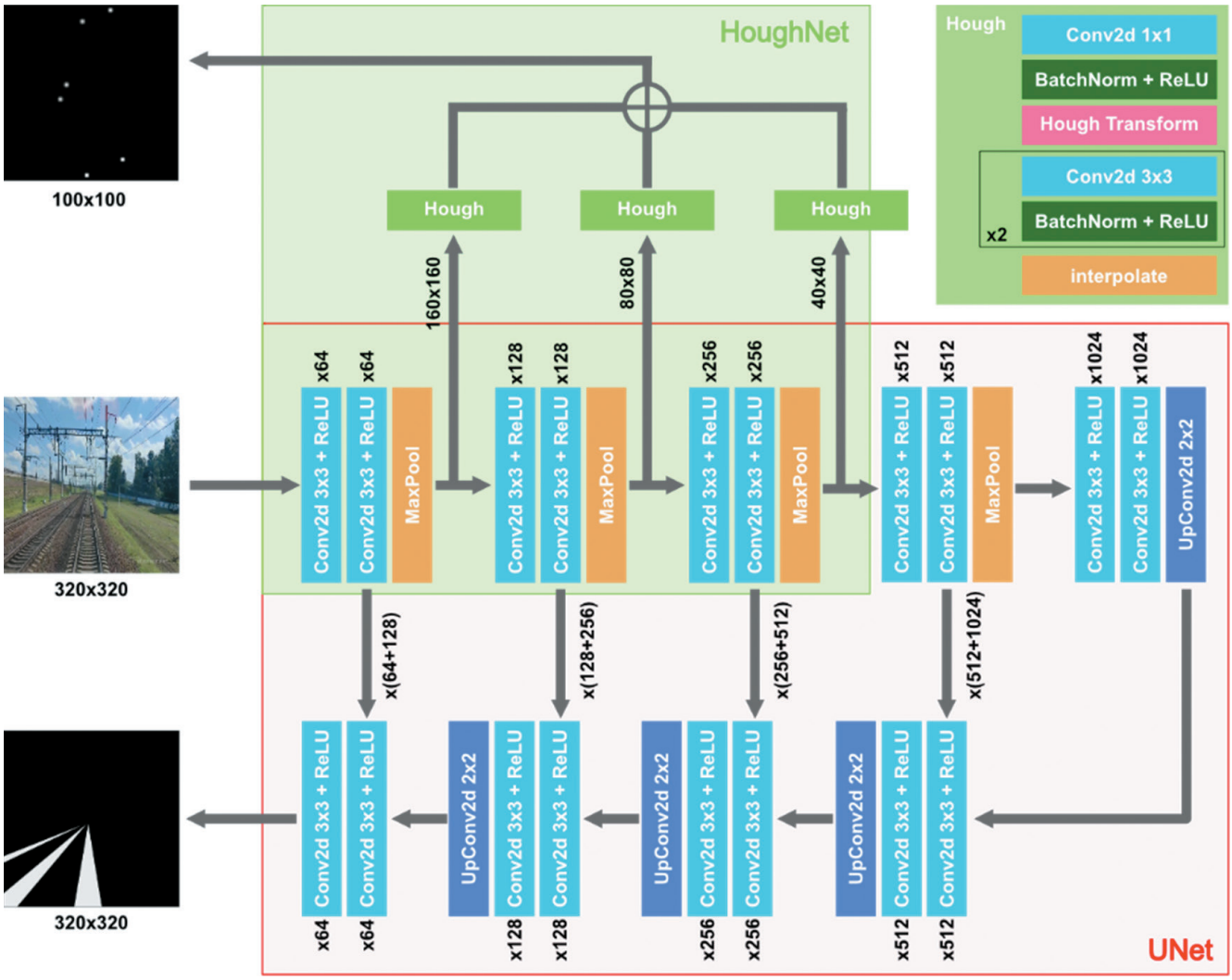

Рис. 8. Структурная схема используемых предиктивных моделей
(красным блоком выделена структура сверточной сети UNet, зеленым – сеть HoughNet)

как «3-ий путь» – линии с углом наклона, близким к 180 градусам. Графа «ошибки первого рода» отражает наличие ложноположительной ошибки предсказания, «ошибки второго рода» – наличие ложноотрицательного предсказания. Графа «FPS» отражает среднее число обрабатываемых предиктивной моделью изображений за одну секунду времени.

По результатам эксперимента было установлено, что использование предлагаемого метода в составе предиктивной модели сопоставимо по точности с классическим методом семантической сегментации изображений. Использование нелинейного интервала квантования параметрического пространства Хафа прогнозируемо повышает точность предсказания модели в рамках зоны интереса и, соответственно, понижет точность вне такой зоны. Разработанная для тестирования модель сети HoughNet обладает большей производительностью на 42% при одинаковых аппаратных ресурсах, однако стоит отметить наличие ошибок второго рода. Следовательно, при использовании более сложных классификаторов в модели для определения карты признаков, выигрыш в производительности может быть ниже или полностью нивелирован.

К преимуществам предлагаемого метода следует отнести однозначное математическое определение прямой линии, благодаря чему появляются возможности экстраполяции на прямом профиле пути. Это обеспечивает устойчивость к определению границы пути при естественных атмосферных искажениях для устойчивых (туман, дым) и динамических (снег, дождь) погодных условий, а также в случаях, когда часть пути перекрыта каким-либо объектом.

Заключение

В рамках данной работы был предложен и рассмотрен альтернативный метод определения границ железнодорожных путей на изображении с помощью геометрических примитивов в виде прямых линий. Предложенный метод построен на вычислительном алгоритме преобразования Хафа.

Для уменьшения времени построения Хаф-образа анализируемого изображения было предложено два

Табл. 2. Результат работы предиктивных моделей

	mIoU (1-ый путь)	mIoU (2-ой путь)	mIoU (3-ий путь)	Ошибки первого рода	Ошибки второго рода	FPS
UNet	92,86%	93,10%	94,54%	–	–	~12
HoughNet (линейный) Θ=100, *R*=100	95,31%	95%	97,40%	–	+	~17
HoughNet (нелинейный) Θ=100, *R*=100, γ=1	95,53%	92,72%	89,33%	–	+	~17
HoughNet (нелинейный) Θ=100, *R*=100, γ=10	98,01%	83,07%	72,10%	–	–	~17

варианта компоновки параметрического пространства в дискретной форме: линейное и нелинейное квантования. Предложенные варианты компоновки пространства учитывают характер работы предиктивной модели комплекса АПК ФИД, а также его физико-технические характеристики, в виде расположения и ориентирования видеокамер на железнодорожный станции. Введено понятие зоны интереса, которая определяется нелинейным распределением для угла наклона прямой линии, описывающей границу железнодорожного пути. В ходе исследования был произведен сравнительный анализ предложенных вариантов компоновки, в рамках которого были рассчитаны ошибки интерполяции.

Для оценки применимости метода к решаемой задаче, качества и скорости его работы в составе предиктивной модели была разработана сверточная сеть HoughNet. В результате был проведен сравнительный анализ разработанной сети с моделью семантической сегментации изображений UNet. Точность работы сети HoughNet оказалась сопоставимой с методом решения задачи с помощью семантической сегментации, однако использование нелинейного интервала квантования Хаф-образа позволило повысить точность прогнозирования в заданной зоне интереса. Модель сети HoughNet оказалась на 42% быстрее модели UNet.

Таким образом, предложенный метод представляет альтернативное решение для технической обеспечивающей задачи комплекса АПК ФИД по определению границ железнодорожного пути и может быть применен на вычислительном графе предиктивной модели комплекса АПК ФИД.

Библиографический список

1. LeCun Y., Boser B., Denker J.S. et al. Backpropagation Applied to Handwritten Zip Code Recognition // Neural Computation. 1989. Vol. 1(4). Pp. 541-551.

2. Zhao H., Shi J., Qi X. et al. Pyramid scene parsing network // Proceedings of the IEEE conference on computer vision and pattern Recognition (CVPR). 2017. Jun. Pp. 2881-2890.

3. He K., Zhang X., Ren S. et al. Deep residual learning for image recognition // Proceedings of the IEEE conference on computer vision and pattern Recognition (CVPR). 2016. Jun. Pp. 770-778.

4. Chen L.-C., Zhu Y., Papandreou G. et al. Encoder-decoder with atrous separable convolution for semantic image segmentation // European Conference on Computer Vision (ECCV). 2016. Sep. Pp. 833-851.

5. Chollet F. Xception: Deep learning with depthwise separable convolutions // Proceedings of the IEEE conference on computer vision and pattern Recognition (CVPR). 2017. Jun. Pp. 1251-1258.

6. Chen L.-C., Papandreou G., Kokkinos I. et al. Deeplab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2018. Vol. 40(4). Pp. 834-848.

7. Hernández D., Cecilia J.M., Cano J.C. et al. Flood detection using real-time image segmentation from unmanned aerial vehicles on edge-computing platform // Remote Sensing. 2022. Vol. 14(1). P. 223.

8. Ronneberger O., Fischer P., Brox T. U-net: Convolutional networks for biomedical image segmentation // Medical Image Computing and Computer-Assisted Intervention (MICCAI) 2015: 18th International Conference. Munich: Springer International Publishing, 2015. Pp. 234-241.

9. Hough P.V. Method and means for recognizing complex patterns. US Patent No. 3,069,654 (1962).

10. Brady M.L., Yong W. Fast Parallel Discrete Approximation Algorithms for the Radon Transform // Proceedings of the fourth annual ACM symposium on Parallel algorithms and architectures (SPAA). 1992. Pp. 91-99.

11. Canny J. A computational approach to edge detection // IEEE Transactions on Pattern Analysis and Machine Intelligence. 1986. Vol. 8. Pp. 679-698.

12. Sobel I. History and Definition of the Sobel Operator. 2014. URL: https://www.researchgate.net/publication/239398674_An_Isotropic_3x3_Image_Gradient_Operator (дата обращения: 24.06.2024).

13. Bradley D., Roth G. Adaptive thresholding using the integral image // Journal of graphics tools. 2007. Vol. 12(2). Pp. 13-21.

14. Cheng C., Kalloor J. Analysis of High Level implementations for Recursive Methods on GPUs. URL: https://people.eecs.berkeley.edu/~kubitron/courses/cs262a-F21/projects/reports/project16_report_ver3.pdf (дата обращения: 24.06.2024).

15. Borgefors G. Distance transformations in digital images // Computer vision, graphics, and image processing. 1986. Vol. 34(3). Pp. 344-371.

16. Murphy A.H. The Finley Affair: A Signal Event in the History of Forecast Verification // Weather and Forecasting. 1996. Vol. 11(1). Pp. 3-20.

17. Lee J.-T., Kim H.-U., Lee C. et al. Semantic line detection and its applications // The International Conference on Computer Vision (ICCV). 2017. Pp. 3229-3237.

18. Гольденберг Л.М., Матюшкин Б.Д. Цифровая обработка сигналов. М.: Радио и связь, 1985. 312 с.

References

1. LeCun Y., Boser B., Denker J.S. et al. Backpropagation Applied to Handwritten Zip Code Recognition. *Neural Computation* 1989;1(4):541-551.

2. Zhao H., Shi J., Qi X. et al. Pyramid scene parsing network. In: Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR); 2017. Pp. 2881-2890.

3. He K., Zhang X., Ren S. et al. Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR); 2016. Pp. 770-778.

4. Chen L.-C., Zhu Y., Papandreou G. et al. Encoder-decoder with atrous separable convolution for semantic image segmentation. In: European Conference on Computer Vision (ECCV); 2016. Pp. 833-851.

5. Chollet F. Xception: Deep learning with depthwise separable convolutions. In: Proceedings of the IEEE conference on computer vision and pattern Recognition (CVPR); 2017. Pp. 1251-1258.

6. Chen L.-C., Papandreou G., Kokkinos I. et al. Deeplab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 2018;40(4):834-848.

7. Hernández D., Cecilia J.M., Cano J.C. et al. Flood detection using real-time image segmentation from unmanned aerial vehicles on edge-computing platform. *Remote Sensing* 2022;14(1):223.

8. Ronneberger O., Fischer P., Brox T. U-net: Convolutional networks for biomedical image segmentation. In: Medical Image Computing and Computer-Assisted Intervention (MICCAI) 2015: 18th International Conference. Munich: Springer International Publishing; 2015. Pp. 234-241.

9. Hough P.V. Method and means for recognizing complex patterns. US Patent No. 3,069,654 (1962).

10. Brady M.L., Yong W. Fast Parallel Discrete Approximation Algorithms for the Radon Transform. In: Proceedings of the fourth annual ACM symposium on Parallel algorithms and architectures (SPAA); 1992. Pp. 91-99.

11. Canny J. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 1986;8:679-698.

12. Sobel I. History and Definition of the Sobel Operator. (accessed 24.06.2024). Available at: https://www.researchgate.net/publication/239398674_An_Isotropic_3x3_Image_Gradient_Operator.

13. Bradley D., Roth G. Adaptive thresholding using the integral image. *Journal of graphics tools* 2007;12(2):13-21.

14. Cheng C., Kalloor J. Analysis of High Level implementations for Recursive Methods on GPUs. (accessed 24.06.2024). Available at: https://people.eecs.berkeley.edu/~kubitron/courses/cs262a-F21/projects/reports/project16_report_ver3.pdf.

15. Borgefors G. Distance transformations in digital images. *Computer vision, graphics, and image processing* 1986;34(3):344-371.

16. Murphy A.H. The Finley Affair: A Signal Event in the History of Forecast Verification. *Weather and Forecasting* 1996;11(1):3-20.

17. Lee J.-T., Kim H.-U., Lee C. et al. Semantic line detection and its applications. In: The International Conference on Computer Vision (ICCV); 2017. Pp. 3229-3237.

18. Goldenberg L.M., Matyushkin B.D. [Digital signal processing]. Moscow: Radio i sviaz; 1985. (in Russ.)

Сведения об авторе

Полевский Илья Сергеевич – главный эксперт, Акционерное общество «Научно-исследовательский и проектно-конструкторский институт информатизации, автоматизации и связи на железнодорожном транспорте» (АО «НИИАС»), Нижегородская улица, дом 27, строение 1, г. Москва, Российская Федерация, 109029, e-mail: i.polevskiy@vniias.ru

About the author

Ilia S. Polevskiy, Chief Expert, Joint Stock Company Research and Design Institute for Information Technology, Signalling and Telecommunications in Railway Transportation (JSC NIIAS), 27, bldg 1 Nizhegorodskaya St., Moscow, 109029, Russian Federation, e-mail: i.polevskiy@vniias.ru.

Вклад автора в статью

Полевский И.С. выполнил обзор библиографических источников, осуществил постановку и формализацию задачи, выполнил разработку метода определения границы железнодорожного пути; ввел определение зоны интереса и предложил варианты дискретизации параметрического пространства предложенного метода, а также привел для них численные оценки качества работы; разработал программное обеспечение аннотации данных и оценочную предиктивную модель; выполнил постановку эксперимента и провел сравнительный анализ результатов.

Конфликт интересов

Автор заявляет об отсутствии конфликта интересов.