

УДК 004.932.2:004.896

***МЕТОД ВИЗУАЛЬНОЙ ОЦЕНКИ НАКЛОННОЙ ПОВЕРХНОСТИ
МОБИЛЬНЫМ РОБОТОМ НА ОСНОВЕ НЕЙРОСЕТЕВОЙ ОБРАБОТКИ
ИЗОБРАЖЕНИЙ***

Кушнирский Я.А.¹

Магистрант

Московский технический университет связи и информатики

Россия, Москва

Аннотация. В статье рассматривается метод визуальной оценки наклонной поверхности мобильным роботом на основе нейросетевой обработки изображений. Предложен модуль, включающий детектор пандуса, регрессор угла наклона и регрессор координат четырех ключевых точек видимой области пандуса. Для обучения и проверки моделей сформирован набор изображений, полученный с камеры робота на стенде с пандусами углов 5–30°. Экспериментально показано, что детектор достигает точности классификации 95,8 %, регрессор угла – средней абсолютной ошибки 2,03°, а регрессор ключевых точек – средней абсолютной ошибки 11,4 пикселя. Полученные параметры могут использоваться для последующей оценки положения робота перед заездом.

Ключевые слова: мобильный робот, компьютерное зрение, наклонная поверхность, пандус, MobileNetV2, регрессия угла, ключевые точки, визуальная навигация.

***METHOD FOR VISUAL ESTIMATION OF AN INCLINED SURFACE BY A
MOBILE ROBOT BASED ON NEURAL IMAGE PROCESSING***

¹ Научный руководитель – к.т.н., доцент Верба В.А.

Kushnirskiy Y.A.

Graduate Student

Moscow Technical University of Communications and Informatics

Russia, Moscow

Abstract. The article considers a method for visual estimation of an inclined surface by a mobile robot based on neural image processing. A module is proposed that includes a ramp detector, an inclination angle regressor, and a regressor for the coordinates of four key points of the visible ramp area. For training and testing the models, an image dataset was collected from the robot camera on a test bench with ramps of α . The experiments showed that the detector achieved a classification accuracy of 95.8 %, the angle regressor achieved a mean absolute error of θ , and the keypoint regressor achieved a mean absolute error of 11.4 pixels. The obtained parameters can be used for subsequent estimation of the robot position before ramp traversal.

Keywords: mobile robot, computer vision, inclined surface, ramp, MobileNetV2, angle regression, keypoints, visual navigation.

Введение

Автономные мобильные роботы при движении в помещениях и на промышленных объектах сталкиваются не только с препятствиями на горизонтальной плоскости, но и с наклонными участками: пандусами, рампами и переходами между зонами разной высоты. Для корректного прохождения таких участков робот должен предварительно определить наличие наклонной поверхности и оценить ее основные параметры. В отличие от обычного препятствия, пандус может рассматриваться как проходимый элемент маршрута, если положение робота и характеристики платформы позволяют выполнить заезд.

В исследованиях проходимости мобильных роботов подчеркивается, что возможность движения зависит от геометрических и физических свойств поверхности, включая уклон, форму участка и характер покрытия [1]. Применение лидаров, стереокамер и плотных карт глубины повышает информативность системы, но усложняет аппаратную часть. Для малогабаритных платформ более рациональным представляется подход, основанный на фронтальной монокулярной камере и вычислительно компактных нейросетевых моделях; сходная логика используется в низкобюджетных робототехнических системах, где навигационные задачи решаются без лидара и отдельного графического ускорителя [2]. В современных работах по самообучаемой оценке проходимости также выделяются уклон поверхности, семантические признаки и особенности конкретной платформы как ключевые факторы безопасного движения [9].

Цель и задачи исследования

Цель статьи состоит в разработке и экспериментальной оценке метода визуальной оценки наклонной поверхности мобильным роботом по изображению с фронтальной монокулярной камеры. Для достижения цели формализуются выходные параметры визуального модуля, описываются набор данных и структура моделей, а также оценивается точность детектора пандуса, регрессора угла наклона и регрессора ключевых точек.

Для достижения цели решаются задачи: определить выходные параметры визуального модуля; описать набор данных и структуру нейросетевых моделей; оценить точность обнаружения пандуса, регрессии угла и предсказания ключевых точек; обозначить ограничения применения метода.

Основная часть

Наклонная поверхность отличается от произвольного препятствия тем, что может быть проходимым элементом маршрута. Поэтому мобильному роботу недостаточно только обнаружить пандус в кадре: необходимо получить

параметры, пригодные для последующей оценки положения перед заездом. В рассматриваемой постановке к таким параметрам относятся вероятность наличия пандуса, угол наклона и координаты характерных точек его видимой области.

Входом визуального модуля является RGB-изображение, получаемое с фронтальной камеры робота. На выходе формируются вероятность наличия пандуса p_{ramp} , оценка угла наклона α и координаты четырех ключевых точек: BL , BR , TL , TR . Точки BL и BR соответствуют левому и правому нижним углам въезда, а TL и TR – левому и правому верхним углам видимой наклонной области.

Формально результат работы модуля можно представить следующим образом:

$$I \rightarrow \{p_{\text{ramp}}, \alpha, P\}, P = \{BL, BR, TL, TR\}$$

где I – входное RGB-изображение; p_{ramp} – вероятность наличия пандуса в кадре; α – оценка угла наклона; P – множество ключевых точек видимой области пандуса.

Такое представление не восстанавливает полную трехмерную геометрию сцены, но дает компактное описание наклонной поверхности, пригодное для последующей проверки центрирования, разворота и дистанции до въезда. Конкретные правила позиционирования зависят от ширины шасси, типа привода, высоты установки камеры и допустимых маневров робота, поэтому в данной статье подробно не рассматриваются.

Задача разделена на три подзадачи: обнаружение пандуса, оценку угла наклона и определение ключевых точек. Такая декомпозиция позволяет локализовать ошибку на уровне классификации, угловой регрессии или предсказания координат. Визуальная оценка проходимости может выполняться без явного восстановления глубины, если признаки изображения используются для выделения проходимых и непроходимых областей [3]. Для пандуса

существенны не только признаки объекта, но и его границы, что согласуется с направлением легких моделей, совмещающих анализ семантики и контуров [4].

В качестве базовой архитектуры выбрана MobileNetV2, ориентированная на устройства с ограниченными вычислительными ресурсами [5]. Это позволяет использовать визуальный модуль на малогабаритной платформе без отдельной графической станции.

Для обучения и проверки моделей был изготовлен испытательный стенд из шести пандусов с углами 5° , 10° , 15° , 20° , 25° и 30° . Съемка выполнялась с камеры мобильного робота при разных расстояниях до пандуса и небольших боковых отклонениях. Для обучения детектора дополнительно использовались негативные примеры: изображения пола, стен, мебели и других объектов без пандуса. Распределение изображений по углам наклона и классу «нет пандуса» представлено в таблице 1.

Таблица 1. Состав набора данных

Угол наклона пандуса	Количество изображений	Доля в датасете
0°	420	20.4%
5°	280	13.6%
10°	310	15.1%
15°	340	16.5%
20°	290	14.1%
25°	240	11.7%
30°	180	8.7%

Изображения приводились к размеру 224×224 пикселя. Набор данных разделялся на обучающую, валидационную и тестовую выборки в пропорции 80/10/10 с сохранением пропорций классов. Для детектора использовалась бинарная постановка: «пандус присутствует» или «пандус отсутствует». Для регрессора угла учитывались только изображения с пандусом, поскольку для класса «нет пандуса» угол наклона не имеет физического смысла.

Разметка ключевых точек выполнялась вручную: для каждого изображения с пандусом фиксировались координаты BL , BR , TL , TR . Качество такой разметки существенно влияет на устойчивость обучения моделей компьютерного зрения [6], а сама логика аннотирования соответствует подходу, при котором объект описывается через визуально выделяемые границы и характерные области [7]. Визуальный модуль включает три модели на общей архитектурной основе: детектор пандуса, регрессор угла наклона и регрессор ключевых точек. Третья модель предсказывает координаты в нормализованном координатном диапазоне $[0; 1]$, что снижает зависимость выхода модели от исходного разрешения изображения. Для обучения использовался оптимизатор Adam, применяемый в нейросетевых моделях благодаря адаптивной оценке моментов градиента [8]. Примеры изображений, использованных при формировании набора данных, приведены на рисунке 1

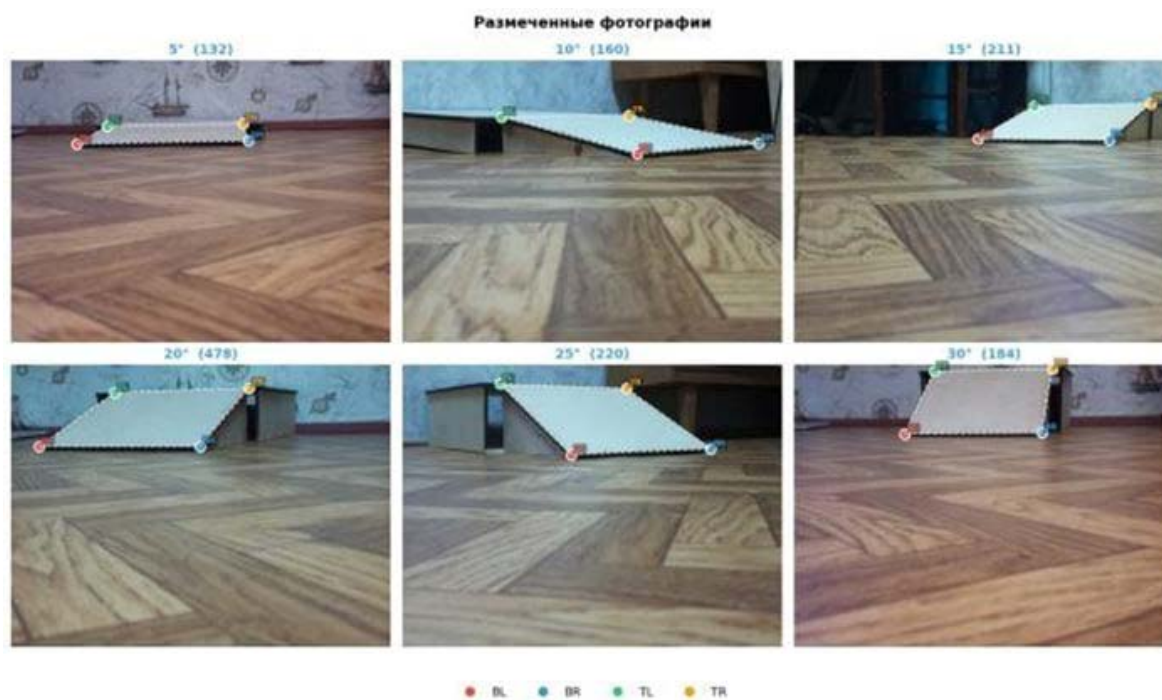


Рисунок 1. Примеры изображений пандусов, полученных с камеры мобильного робота

Примечание – разработано автором.

Обобщенная последовательность обработки изображения визуальным модулем представлена на рисунке 2.

Использование трех моделей вместо единой схемы прямого формирования управляющего действия делает результат более интерпретируемым. Детектор, регрессор угла и регрессор ключевых точек можно оценивать независимо, а их ошибки – учитывать при настройке последующих правил позиционирования.



Рисунок 2. Последовательность нейросетевой обработки изображения пандуса

Примечание – разработано автором.

Результаты

Для оценки качества детектора использовались следующие метрики: доля правильных классификаций (accuracy), точность положительного класса (precision), полнота (recall), F1-мера и площадь под ROC-кривой (AUC-ROC). Для регрессионных моделей применялись средняя абсолютная ошибка (Mean

Absolute Error, MAE), среднеквадратическая ошибка (Root Mean Squared Error, RMSE), коэффициент детерминации R^2 , а также доля корректно предсказанных ключевых точек в пределах заданного допуска (Percentage of Correct Keypoints, PCK). Результаты оценки детектора пандуса и регрессора угла наклона приведены в таблице 2.

Таблица 2. Результаты детектора и регрессора угла

Модель	Метрика	Тестовая выборка
Детектор пандуса	Accuracy	95,8 %
Детектор пандуса	Precision	95,3 %
Детектор пандуса	Recall	96,1 %
Детектор пандуса	F1-score	95,7 %
Детектор пандуса	AUC-ROC	98,9 %
Регрессор угла	MAE	2,03°
Регрессор угла	RMSE	2,58°
Регрессор угла	R^2	0,965
Регрессор угла	Max error	7,1

Регрессор угла показал среднюю абсолютную ошибку 2,03° на тестовой выборке. При дискретности испытательного стенда 5° такая точность достаточна для предварительной оценки категории уклона. Максимальная ошибка 7,1° указывает, что отдельные кадры могут давать заметное отклонение, особенно при неблагоприятном ракурсе или совпадении границ пандуса с фоном. Поэтому при последующей интеграции в контур движения целесообразно использовать не единичное предсказание, а сглаживание оценки по нескольким последовательным кадрам.

Регрессор ключевых точек оценивался по MSE нормализованных координат, средней абсолютной ошибке в пикселях и PCK@0.05. На тестовой выборке средняя ошибка составила 11,4 пикселя при разрешении 640 × 480, а PCK@0.05 – 91,7 %. Такой уровень точности достаточен для грубой

Дневник науки | www.dnevniknauki.ru | СМИ ЭЛ № ФС 77-68405 ISSN 2541-8327

геометрической параметризации пандуса: оценки положения нижней границы въезда, видимой ширины и перспективного смещения. Результаты оценки регрессора ключевых точек представлены в таблице 3.

Таблица 3. Результаты регрессора ключевых точек

Метрика	Обучающая выборка	Тестовая выборка
MSE нормализованных координат	0,0031	0,0048
MAE при 640 × 480, пиксели	8,2	11,4
PCK@0.05	94,3 %	91,7 %

Данные таблицы 3 показывают, что ошибка регрессора ключевых точек не предназначена для точного метрического восстановления сцены, однако достаточна для первичной геометрической параметризации пандуса. Координаты точек могут использоваться как признаки положения видимой области: смещения въезда в кадре, относительной ширины нижней границы и перспективного сужения. Эти признаки пригодны для последующей проверки центрирования, разворота робота и допустимой дистанции до въезда.

Полученные результаты подтверждают работоспособность визуального модуля в условиях испытательного стенда, но область применимости метода ограничена фиксированной геометрией пандусов и контролируемым освещением. Для переноса метода на реальные пандусы требуется расширить выборку за счет других материалов, покрытий, высот установки камеры, теней, бликов и частичных перекрытий, а также разделять данные по сериям съемки или сценам, чтобы близкие кадры одного ракурса не попадали одновременно в обучающую и тестовую выборки. Отдельные исследования показывают, что перенос визуальных моделей проходимости на реальные сцены требует учета доменного сдвига между синтетическими и натурными данными [10].

Заключение

Разработан метод визуальной оценки наклонной поверхности мобильным роботом на основе нейросетевой обработки изображений, ориентированный на

Дневник науки | www.dnevniknauki.ru | СМИ ЭЛ № ФС 77-68405 ISSN 2541-8327

использование фронтальной монокулярной камеры и вычислительно компактной архитектуры MobileNetV2. Экспериментальная проверка на наборе изображений пандусов с углами 5° – 30° показала работоспособность предложенного визуального модуля: детектор достиг ассурасы 95,8 %, регрессор угла – MAE 2,03°, а регрессор ключевых точек – MAE 11,4 пикселя при разрешении 640×480 .

Полученные результаты показывают, что наклонная поверхность может быть представлена компактным набором визуальных параметров: фактом наличия пандуса, углом наклона и координатами четырех характерных точек. Данный результат образует входной уровень для последующего позиционирования робота перед заездом; дальнейшее развитие работы целесообразно связать с построением геометрических правил и их адаптацией к различным конфигурациям мобильной платформы.

Библиографический список

1. Sevastopoulos C., Konstantopoulos S. A Survey of Traversability Estimation for Mobile Robots : arXiv:2204.10883. – 2022. – DOI: 10.48550/arXiv.2204.10883.
2. Kim T., Lim S., Shin G., Sim G., Yun D. An Open-Source Low-Cost Mobile Robot System With an RGB-D Camera and Efficient Real-Time Navigation Algorithm // IEEE Access. – 2022. – Vol. 10. – P. 127871–127881. – DOI: 10.1109/ACCESS.2022.3226784.
3. Frey J., Mattamala M., Chebrolu N., Cadena C., Fallon M., Hutter M. Fast Traversability Estimation for Wild Visual Navigation : arXiv:2305.08510. – 2023. – DOI: 10.48550/arXiv.2305.08510.
4. Liao Y., Kang S., Li J., Liu Y., Liu Y., Dong Z., Yang B., Chen X. Mobile-Seed: Joint Semantic Segmentation and Boundary Detection for Mobile Robots : arXiv:2311.12651. – 2023. – DOI: 10.48550/arXiv.2311.12651.

5. Sandler M., Howard A., Zhu M., Zhmoginov A., Chen L.-C. MobileNetV2: Inverted Residuals and Linear Bottlenecks // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. – 2018. – P. 4510–4520.
6. Sager C., Janiesch C., Zschech P. A Survey of Image Labelling for Computer Vision Applications : arXiv:2104.08885. – 2021. – DOI: 10.48550/arXiv.2104.08885.
7. Russell B. C., Torralba A., Murphy K. P., Freeman W. T. LabelMe: A Database and Web-Based Tool for Image Annotation // International Journal of Computer Vision. – 2008. – Vol. 77. – P. 157–173. – DOI: 10.1007/s11263-007-0090-8.
8. Kingma D. P., Ba J. Adam: A Method for Stochastic Optimization // Proceedings of the 3rd International Conference on Learning Representations. – 2015.
9. Jeon Y., Son E. I., Seo S.-W. Follow the Footprints: Self-supervised Traversability Estimation for Off-road Vehicle Navigation based on Geometric and Visual Cues // Proceedings of the IEEE International Conference on Robotics and Automation. – 2024. – arXiv:2402.15363. – DOI: 10.48550/arXiv.2402.15363.
10. Vecchio G. et al. Terrain traversability prediction through self-supervised learning and unsupervised domain adaptation on synthetic data // Autonomous Robots. – 2024. – Vol. 48. – Article 4. – DOI: 10.1007/s10514-024-10158-4.