

Сравнение методов решения задачи определения пиков интенсивности на эхолокационной трассе с помощью алгоритмов машинного обучения

И. С. Писарев, А. А. Ухов

Санкт-Петербургский государственный электротехнический университет
«ЛЭТИ» им. В.И. Ульянова (Ленина)

ilya.pisarev91@gmail.com, aauhov@yandex.ru

Аннотация. В рамках данного исследования проведено сравнительное изучение методов обработки данных для обнаружения объектов на эхолокационной трассе с использованием алгоритмов машинного обучения. Актуальность работы связана с необходимостью автоматизации анализа данных в сферах дистанционного зондирования, экологического мониторинга и систем компьютерного зрения. Методология включает сбор данных, полученных при регистрации интенсивности лазерного излучения в разнообразных условиях, и их первичную обработку. В работе предложен подход, основанный на анализе всей эхолокационной трассы с прогнозированием расстояния до объекта, а также на выделении ключевых объектов с последующей классификацией. Для обучения моделей использовался специально подготовленный датасет с аннотированными графиками, а оценка эффективности проводилась с применением метрик RMSE (для задачи регрессии) и Accuracy, F1-score (для классификации).

Эксперименты выявили, что наиболее результативным является метод сегментации эхолокационной трассы с последующей классификацией выделенных областей. Алгоритм градиентного бустинга показал наивысшую эффективность, достигая значений Accuracy и F1-score 98%. Предложенный подход демонстрирует высокую эффективность и потенциал для внедрения в системы автоматизированного анализа лазерных сигналов, а также для адаптации к решению смежных задач, таких как фильтрация помех и оптимизация обработки данных в реальном времени.

Ключевые слова: обработка сигналов; методы разметки данных; машинное обучение

I. ВВЕДЕНИЕ

При решении задач измерения расстояния между двумя объектами часто используют лазерные дальномеры. Высокая точность измерений, возможность измерения больших дистанций, скорость измерений, простота интеграции измерителя в общее программное обеспечение позволяют применять их в геодезии, навигации, картографии и других областях.

При измерении дистанций с помощью лазерных дальномеров на большие расстояния могут возникать проблемы, связанные с точностью при эксплуатации прибора в экстремальных погодных условиях, таких как дождь, снег, туман. Помимо погодных помех на пути луча могут встречаться препятствия, которые могут быть восприняты устройством обработки сигнала как ложные объекты. Поэтому использование новых программных алгоритмов обработки сигнала, модернизация и переработка существующих аппаратных решений

востребованы и позволяют улучшить точность устройства, а также повысить скорость.

A. Методы измерения временного промежутка

Существует множество алгоритмов измерения расстояния импульсных лазерных дальномеров [1]. Широкое распространение получил времяпролетный метод. Он основан на измерении временного промежутка, за который излучение проходит расстояние до объекта, отражается от него и возвращается назад.

Данный метод требует точного определения старта излучения и времени, когда отраженный сигнал вернулся обратно. Временная точка (момент), когда должна произойти генерация сигналов запуска и остановки, является результатом наблюдения различных характерных точек на форме волны принимаемых сигналов [2]. Существует множество методов измерения для выполнения этой задачи, которые можно разделить на три основные группы в зависимости от подхода к определению момента остановки: обнаружения пика, постоянное пороговое обнаружение, адаптивное пороговое обнаружение.

Постоянное пороговое обнаружение с распознаванием по переднему краю один из самых просто реализуемых аппаратных решений. Смысл метода заключается в генерации сигналов запуска и остановки по установленному заранее уровню напряжения. Основным недостатком метода является то, что момент генерации сигнала остановки и запуска зависит от изменения мощности принимаемого импульса лазерного излучения, как показано на рис. 1 (t_1 и t_2):

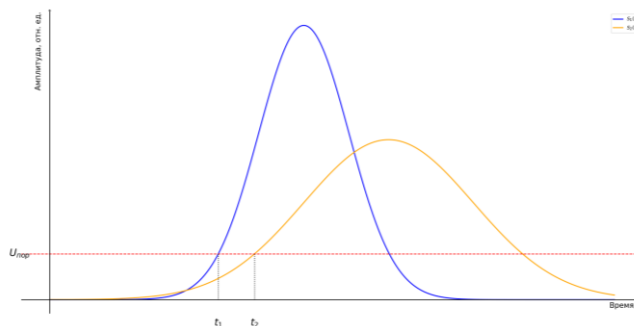


Рис. 1. Характеристическая кривая излучения

Целью более сложных методов является обеспечение независимости измерения стоп-сигнала от изменения амплитуды принимаемого импульса. Они основаны на привязке генерации сигнала остановки к определению

пикового напряжения. Для цифровых сигналов методы аналогичны, однако дискретность и квантование сигналов снижают точность. Для её повышения требуются либо высокая частота дискретизации, либо сверхкороткие импульсы, что усложняет аппаратуру.

В. Проведение измерений

При проектировании лазерных дальномеров необходимо задать тип объекта, до которого будут производиться измерения, а также учитывать типовые виды объектов, которые могут встречаться на пути зондирующего луча.

Классификация осуществляется на основе соотношений размеров объекта с размерами рабочего поля дальномера. В основном выделяют следующие типы:

- Крупноразмерная цель — объект, угловые размеры которого превышают рабочее поле дальномера, образуемое областью пространства, охватываемого излучением дальномера и зоной, где дальномер может принимать сигнал. Соответственно малоразмерная цель не превышает рабочее поле.
- Основная цель — объект с заданными габаритами и коэффициентом яркости, для измерения дальности которого предназначен прибор. В качестве основной цели обычно берут объект 2х3 м коэффициентом яркости 0.2 [3].

При проведении зондирования на пути следования пучка могут встречаться достаточно распространенные помехи типа «куст», «дерево», «пылевое облако», а также помеха обратного рассеяния, вызванная отражением лазерного зондирующего пучка от аэрозоля. Схема горизонтальной эхолокационной трассы представлена на рис. 2.

В дальнометрии для борьбы с типовыми помехами применяют адаптивные алгоритмы, позволяющие устанавливать динамический порог срабатывания, зависящий от расстояния. Встречаются алгоритмы, анализирующие амплитудные и временные характеристики импульса, отделить полезный сигнал от фоновых искажений. Но помехи, часто неотличимы от настоящего сигнала, что приводит данные алгоритмы к ложным срабатываниям, а в полуавтоматических методах ошибку может вносить субъективность оператора. Поэтому был принят решение разработать алгоритм фильтрации для исключения помех и упрощения проведения измерений для оператора. Идея с анализом обратного рассеяния от зондирующего лазерного пучка применяется в устройствах, анализирующих атмосферу земли [4]. Алгоритмы используют в основном анализ изображений [5–6], в то время как предлагаемый метод основывается только на анализе эхолокационной трассы.

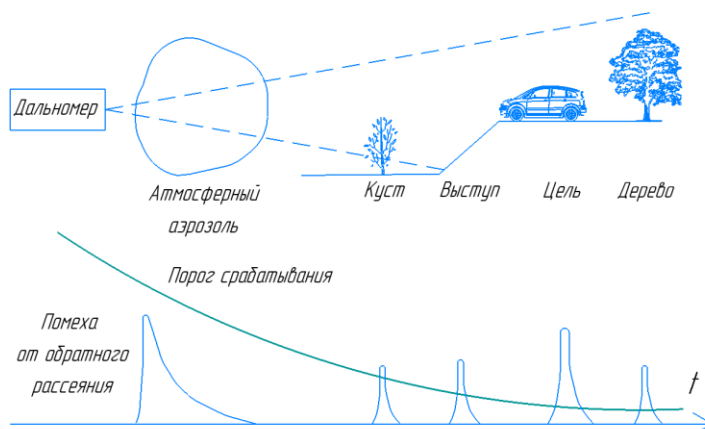


Рис. 2. Схема локационной трассы

II. РАЗРАБОТКА АЛГОРИТМА

Главная сложность при анализе целей на графике зависимости интенсивности лазерного излучения от расстояния заключается в неопределённости множества параметров, необходимых для построения точных математических моделей. Это затрудняет выявление устойчивых закономерностей. Для решения проблемы было выбрано применение методов машинного обучения, так как они позволяют обрабатывать значительные объёмы данных, автоматически выявлять скрытые взаимосвязи между параметрами и формировать прогнозные решения.

При формировании обучающей выборки были выбраны две цели, расположенные на заранее известном расстоянии от точки проведения измерений.

Эхолокационные трассы записывались с использованием одного лазерного дальномера, установленного в фиксированной позиции, но при изменяющихся метеорологических условиях.

Данные в файлах представляют собой набор интенсивностей. Так же в начале файла разработчик указывает параметры, при которых было произведено зондирование:

- частота дискретизации, МГц (необходимо для расчета расстояния);
- коэффициенты усиления.

На рис. 3 представлены визуализированные данные в виде графика.

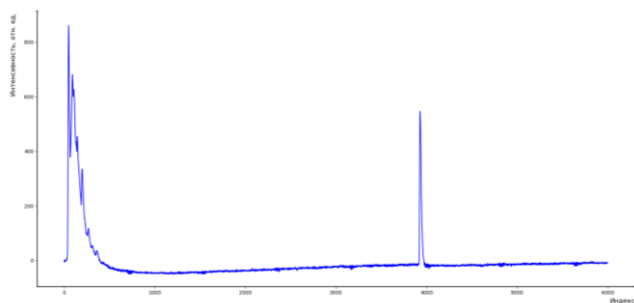


Рис. 3. Визуализация данных эхолокационной трассы

На рисунке отчетливо видно два характерных участка графика: обратное рассеяние от аэрозоля в начале локационной трассы, а затем следует цель. Так же стоит отметить, что для качественного обучения модели выбирались графики, где отсутствуют помехи.

При анализе эхолокационных трасс были разработаны два принципиально различных подхода. Первый подход предполагает предсказание местоположения цели, подразумевая автоматическую фильтрацию возможных помех, встречающихся на пути до объекта. Второй подход направлен на классификацию всех объектов на трассе, включая цели и помехи, с последующим выделением целевого сигнала. Поэтому было принято решение сравнить эффективность этих методов и определить оптимальный для автоматизации анализа лазерных сигналов.

А. Разметка данных

Для каждого из подходов были реализованы разные алгоритмы разметки данных. Для предсказания местоположения расстояния до объекта на эхолокационной трассе выделялся индекс, с которого начинается цель, как это показано на рис. 4.

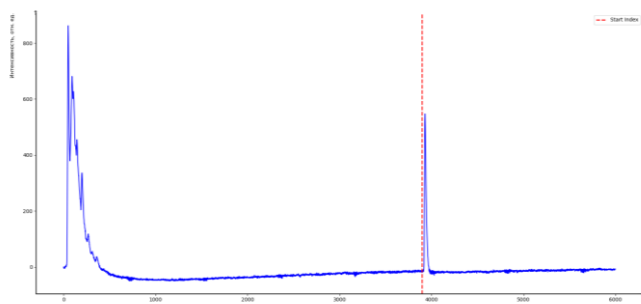


Рис. 4. Разметка эхолокационной трассы для решения задачи регрессии

При обработке графика использовались алгоритмы обнаружения пиков интенсивности и адаптивные пороговые методы.

В рамках второго подхода график делился на сегменты, каждый из которых анализировался отдельно. Алгоритм разметки выделял области, на которых отмечены обратное рассеяние и цель. Так же при подготовке обучающей выборки уравнивалось количество точек каждого кадра, как отдельные параметры записывалась информация о параметрах окружающей среды в момент проведения сбора данных. На рис. 5 представлена визуализация локационной трассы до разметки и после разметки с выделением двух объектов: обратного рассеяния и цели.

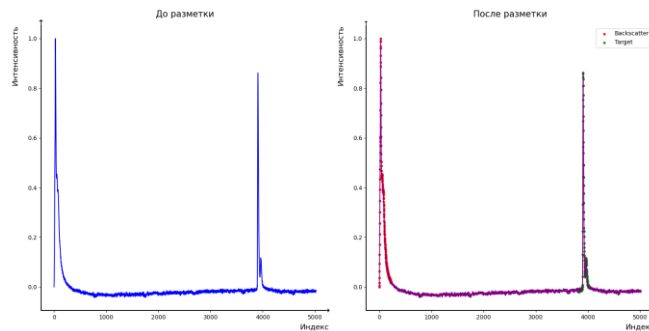


Рис. 5. Схема локационной трассы: а) до разметки, б) после разметки

III. РЕЗУЛЬТАТЫ ЭКСПЕРИМЕНТА

Для оценки эффективности методов машинного обучения в задачах регрессии и классификации, направленных на прогнозирование расстояния до цели и фильтрацию помех, все эксперименты проводились на вычислительной системе на основе процессора AMD Ryzen 7 3700X и оперативной памяти объемом 32 Гб DDR4 (2400 МГц).

Решая задачу регрессии, были выбраны три алгоритма: линейная регрессия, случайный лес и градиентный бустинг. В качестве ключевой метрики использовался RMSE (Root Mean Square Error), который отражает среднеквадратичное отклонение предсказанных значений от фактических. Данная метрика позволяет количественно оценить точность модели в условиях вариативности сигналов и наличия шумов, что критично для задач, требующих высокой достоверности измерений. Для оценки результатов, полученных в ходе эксперимента, пересчитаем полученный результат в расстояние по формуле:

$$L = \frac{n}{2F_{\text{sample}}} c, \quad (1)$$

где L — расстояние [м], F_{sample} — частота дискретизации АЦП [МГц], n — количество отсчетов, c — скорость света [м/с].

Подставив полученное значение целевой метрики в (1), мы можем оценить в метрах, насколько ошибается модель.

ТАБЛИЦА I. СРАВНЕНИЕ МЕТРИК ОЦЕНКИ КАЧЕСТВА ОБУЧЕНИЯ ДЛЯ РАЗНЫХ МОДЕЛЕЙ ПРИ РЕШЕНИИ ЗАДАЧИ РЕГРЕССИИ

Название модели	RMSE, отсчет	Разброс, м	Время идентификации, мкс
Линейная регрессия	113,4	34,2	18,9
Случайный лес	87,67	26,3	126,7
Градиентный бустинг	55,12	15,33	17,9

Из-за большого количества входных параметров модель не смогла найти зависимости. Это ограничило эффективность алгоритмов, и, несмотря на кажущуюся простоту базовых методов, итоговая точность предсказаний оказалась неудовлетворительной. В частности, избыток шумовых признаков и слабая корреляция части параметров с целевой переменной создали дополнительные помехи при обучении.

Для решения задачи классификации объектов на эхолокационной трассе были выбраны три модели: логистическая регрессия, случайный лес и градиентный бустинг. Критериями сравнения стали метрики ассурасу, отражающая общую точность предсказаний, и f1-score, учитывающая баланс между точностью (precision) и полнотой (recall). Эти показатели позволяют комплексно оценить способность моделей не только корректно идентифицировать цели, но и минимизировать ложные срабатывания в условиях помех.

Обучение проводилось на предварительно размеченном датасете, разделенном на обучающую и тестовую выборки, что обеспечило объективность оценки. Условия эксперимента включали анализ графиков интенсивности лазерного излучения с выделением областей обратного рассеяния и целевых объектов. В табл. 2 представлены результаты с полученными метриками оценки качества обученных моделей.

ТАБЛИЦА 1. СРАВНЕНИЕ МЕТРИК ОЦЕНКИ КАЧЕСТВА ОБУЧЕНИЯ ДЛЯ РАЗНЫХ МОДЕЛЕЙ ПРИ РЕШЕНИИ ЗАДАЧИ КЛАССИФИКАЦИИ

Название модели	Accuracy	F1-score	Время идентификации, мкс
Логистическая регрессия	0,67	0,45	4,7
Случайный лес	0,82	0,78	12,7
Градиентный бустинг	0,98	0,98	0,3

Таким образом, лучшие показатели удалось получить при обучении модели градиентного бустинга.

IV. ВЫВОД

В рамках исследования были сопоставлены два подхода к определению пиков интенсивности на эхолокационной трассе. Результаты экспериментов выявили, что метод прогнозирования расстояния до

объекта демонстрирует ограниченную точность из-за неспособности модели выявить устойчивые зависимости в большом количестве входных параметров, характерных для одного объекта. Наиболее результативным оказался подход, основанный на сегментации трассы с последующей классификацией выделенных областей. Для оценки качества классификации использовались метрики Accuracy и F1-score. Сравнение различных моделей машинного обучения показало, что наилучшие результаты достигаются при использовании градиентного бустинга, обеспечивающего Accuracy на уровне 98 % и F1-score 98 %. Таким образом, предложенный подход может быть эффективно применен в задачах автоматического анализа лазерных сигналов, что открывает перспективы его использования в дистанционном зондировании, мониторинге и других областях, связанных с обработкой лазерного излучения.

СПИСОК ЛИТЕРАТУРЫ

- [1] В.Е. Карасик, Лазерные приборы и методы измерения дальности /под ред. В.Е. Карасика. Изд. МГТУ им. Н.Э. Баумана, 2012, 96 с.
- [2] Palojarvi, P. Integrated Electronic and Optoelectronic Circuits and Devices for Pulsed Time-of-Flight Laser Rangefinding; Oulu Press, University of Oulu: Oulu, Finland, 2003.
- [3] В. Вильнер, В. Волобуев, А. Ларюшин, А. Рябокуль, «Достоверность измерений импульсного лазерного Дальномера» // Фотоника. 2013. №3. С. 42–60.
- [4] Romano F., Cimini D., Di Paola F., Gallucci D., Larosa S., Nilo S.T., Ricciardelli E., Lisager B.D., Hutchison K. The Evolution of Meteorological Satellite Cloud Detection Methodologies for Atmospheric Parameter Retrievals. Remote Sens. 2024, 16, 2578. <https://doi.org/10.3390/rs16142578>
- [5] Heidinger A.K., Bearson N., Foster M.J., Li Y., Wanzong S., Ackerman S., Holz R.E., Platnick S., Meyer K. Using Sounder Data to Improve Cirrus Cloud Height Estimation from Satellite Imagers. // J. Atmos. Ocean. Technol. 2019, 36, 1331–1342.
- [6] Weisz E., Baum B.A., Menzel W.P. Fusion of Satellite-Based Imager and Sounder Data to Construct Supplementary High Spatial Resolution Narrowband IR Radiances // JARS 2017, 11, 036022.