

Санкт-Петербургский государственный университет

Артеменко Софья Алексеевна

Выпускная квалификационная работа

Библиотека для работы с данными
дистанционного зондирования земли в
формате GeoTIFF

Уровень образования: бакалавриат

Направление *02.03.03 «Математическое обеспечение и администрирование
информационных систем»*

Основная образовательная программа *СВ.5162.2021 «Технологии программирования»*

Научный руководитель:
старший преподаватель кафедры СП, к. т. н., Ю. В. Литвинов

Рецензент:
Генеральный директор ООО СКЗ, М. Н. Смирнов

Санкт-Петербург
2025

Saint Petersburg State University

Sofya Artemenko

Bachelor's Thesis

A library for working with Earth remote sensing data in GeoTIFF format

Education level: bachelor

Speciality *02.03.03 "Software and Administration of Information Systems"*

Programme *CB.5162.2021 "Programming Technologies"*

Scientific supervisor:
C.Sc., senior lecturer Y.V. Litvinov

Reviewer:
General Director at CVS LLC M.N. Smirnov

Saint Petersburg
2025

Оглавление

Введение	4
1. Постановка задачи	7
2. Обзор существующих аналогов	8
2.1. PROJ	8
2.2. GDAL	8
2.3. libgeotiff	9
2.4. GeographicLib	10
3. Обзор предметной области	11
4. Реализация	14
4.1. Требования	14
4.2. Концепция решения	15
4.3. Тестирование устойчивости проекций к погрешностям . .	17
4.4. Моделирование	18
4.5. Архитектура	21
5. Эксперимент	23
5.1. Условия эксперимента	23
5.2. Исследовательские вопросы	23
5.3. Метрики	23
5.4. Результаты эксперимента	23
Заключение	26
Список литературы	27

Введение

Геолокация на основе изображений [1] — это процесс определения местоположения летательного аппарата на основе анализа изображения, сделанного его камерой. При этом не используется GNSS (Global Navigation Satellite System), либо GNSS используется только для грубой оценки. Такая технология установления местоположения является ключевой в автономной навигации и геоинформационных системах.

Главным этапом геолокации на основе изображений является сопоставление полученного снимка с базой эталонных геопривязанных данных, представленных, например, в формате GeoTIFF.

Файлы GeoTIFF используются для хранения растровых географических изображений с привязанными к ним метаданными о пространственной информации. Они сочетают в себе особенности обычных файлов TIFF (Tagged Image File Format), которые обеспечивают высокое качество и гибкость в представлении изображений, и файлов GIS, которые содержат информацию о географической привязке изображения и их проекционных параметрах. Таким образом, файлы GeoTIFF позволяют хранить растровые изображения с точной географической информацией, что делает их полезными для многих приложений, связанных с географическими данными, таких как дистанционное зондирование, геодезия и картография. Работа с GeoTIFF осложняется большим размером изображений, что не позволяет открывать и хранить их в исходном виде. Ввиду данной особенности GeoTIFF-файл обычно дополнительно делится на тайлы для повышения производительности и допустимости использования в условиях ограниченных ресурсов.

В работе с географическими данными использование широт и долгот зачастую оказывается неудобным, так как эти координаты заданы в сферической системе, где расстояния и углы между точками зависят от их положения на поверхности Земли. Это затрудняет проведение расчетов и визуализацию данных. Для упрощения задачи применяют картографические проекции, которые позволяют преобразовать сферическую систему координат в плоскую. В результате долгота и широта

переводятся в значения x и y , которые можно использовать в привычных расчетах на плоскости, таких как измерение расстояний, построение карт и анализ данных в геоинформационных системах.

Для решения различных задач требуются разные картографические проекции, так как каждая из них обладает своими свойствами, преимуществами и ограничениями. Например, для навигационных целей часто используют проекцию Меркатора, которая сохраняет углы, но искажает площади, особенно вблизи полюсов. Однако ни одна проекция не может полностью избежать погрешностей, так как переход от сферической поверхности к плоскости сопровождается искажениями форм, площадей, углов или расстояний. Выбор подходящей проекции определяется спецификой задачи и допустимыми уровнями искажений в расчетах.

Некоторые исходные предоставляемые GeoTIFF-файлы содержат уже проецированные координаты, поэтому данный фактор необходимо учитывать перед выполнением дальнейших трансформаций и обработок.

Вышеописанные преобразования координат необходимо осуществлять в режиме реального времени, обеспечивая при этом стабильную работу системы даже в условиях ограниченных вычислительных ресурсов. Это требует особого внимания к рациональному использованию памяти и процессорного времени.

Существующие библиотеки предоставляют широкий набор инструментов для работы с геоданными, однако их функциональности часто оказывается недостаточно для решения более специфических и прикладных задач, таких как проецирование больших по площади регионов, требующее использования нескольких наборов параметров для разных частей, проложение траекторий с помощью изображений. Также нуждаются в рассмотрении и более теоретические аспекты работы с ГИС-данными, например, упрощение алгоритмов расчета расстояний. Такие задачи требуют разработки новых методов, которые могут не только повысить точность вычислений, но и сократить затрачиваемые вычислительные ресурсы, что особенно важно для работы с большими

объемами данных и динамическими системами.

Необходимо разработать алгоритм, обеспечивающий согласованность координат в реальном времени с контролируемым уровнем погрешности, при этом учитывающий ограничения по производительности и избегающий избыточных вычислений.

1. Постановка задачи

Целью работы является разработка библиотеки алгоритмов для обеспечения согласованного использования координат, получаемых при работе с изображениями большого размера в формате GeoTIFF. Для ее выполнения были поставлены следующие задачи.

1. Выполнить обзор проблемной области (проекции Земли).
2. Собрать требования к библиотеке.
3. Рассмотреть существующие аналоги.
4. Создать и реализовать алгоритмы.
5. Спроектировать и реализовать библиотеку.
6. Протестировать библиотеку.

2. Обзор существующих аналогов

2.1. PROJ

Библиотека PROJ [6] предоставляет широкий спектр функциональности для работы с географическими координатами и проекциями. Основные возможности PROJ включают: прямое и обратное преобразование координат между географическими системами отсчета, поддержку множества проекций и моделей Земли, проецирование и обратное проецирование, интеграцию с форматами геопространственных данных, а также работу с геодезическими линиями и расстояниями. PROJ предоставляет гибкий API, который поддерживает работу на различных языках программирования, включая C, C++, Python, R и другие. Кроме того, библиотека включает командную утилиту `proj`, позволяющую выполнять геодезические операции напрямую из консоли.

Следует учитывать, что PROJ выполняет достаточно базовые операции по трансформации координат, а также лишние или не относящиеся к нашей задаче преобразования, использующие, например, грид-файлы (точная информация о локальных деформациях) или 3D-геометрию. Поэтому в рамках нашего проекта PROJ может рассматриваться скорее в качестве низкоуровневого компонента, который может быть встроен в более высокоуровневую систему, выполняющую необходимую бизнес-логику и управление параметрами проекций.

Вышеописанные возможности делают PROJ ключевым компонентом для разработки программного обеспечения в области геопространственного анализа и навигации. Распространяется под лицензией, позволяющей свободно использовать и распространять данную библиотеку.

2.2. GDAL

GDAL (Geospatial Data Abstraction Library) [4] — широко используемая библиотека для работы с геопространственными данными. Она поддерживает большое разнообразие растровых и векторных форма-

тов. Предоставляется возможность обрезать растровые изображения и изменять их размер, перепроецировать их в различные системы координат, а также можно составлять из нескольких файлов единый или применять фильтрацию и маскирование. Важно отметить, что применение мозаичного объединения файлов возможно только после приведения составляющих частей к единой системе координат. Библиотека позволяет читать, записывать, преобразовывать и перепроецировать векторные данные.

GDAL, также как и PROJ, не предусматривает механизмов для динамического выбора параметров проекций, отслеживания накопленных искажений, принятия решений о необходимости смены проекции. В связи с этим требуется разработка надстройки, реализующей логику принятия решений и адаптации параметров преобразования в соответствии с текущими условиями.

GDAL является фундаментом для множества ГИС и приложений, таких как QGIS¹, ArcGIS², PostGIS³ и другие. Распространяется под лицензией MIT.

2.3. libgeotiff

libgeotiff [7] — это компактная библиотека на C для работы с GeoTIFF-файлами, позволяющая читать, записывать и редактировать географические метаданные в TIFF-изображениях. Обработка последних осуществляется с помощью libtiff [8]. Она обеспечивает поддержку различных систем координат и проекций, минимизируя зависимости и ресурсозатраты. libgeotiff используется в составе GDAL для обработки геопространственных данных, выполняя функции работы с GeoTIFF на низком уровне, что делает ее важной частью ГИС и картографических приложений.

Данная библиотека может быть особенно полезна, если в ней содер-

¹QGIS — ПО для работы с ГИС, URL: <https://www.qgis.org>, дата обращения: 21.05.2025

²ArcGIS — комплекс геоинформационных программных продуктов компании ESRI, URL: <https://www.esri.com/en-us/arcgis/products/arcgis-online/overview>, дата обращения: 21.05.2025

³PostGIS — расширение для PostgreSQL, поддерживающее работу с географическими данными, URL: <https://postgis.net>, дата обращения: 21.05.2025

жится вся необходимая функциональность для решения конкретной задачи, что позволит избежать проблем с согласованием версий библиотек, используемых в GDAL, и другими задействованными в проекте. Распространяется под лицензией MIT.

2.4. GeographicLib

GeographicLib [5] — это специализированная библиотека для проведения высокоточных геодезических вычислений на поверхности эллипсоида. Она предоставляет алгоритмы для решения как прямой, так и обратной геодезической задачи, а также для расчета расстояний, азимутов и координат между точками с учетом эллипсоидальной формы Земли.

GeographicLib доступна для использования в ряде языков программирования (включая C++, Python), снабжена утилитами командной строки и может интегрироваться в сторонние геопространственные библиотеки, включая PROJ.

Использование GeographicLib в рамках данного проекта не предполагается, однако оно целесообразно для проведения экспериментов и получения достоверных референсных геодезических значений.

Распространяется под лицензией MIT.

3. Обзор предметной области

Существует множество картографических проекций, а именно: эквидистанциальные, эквивалентные и конформные, каждая из которых используется для различных задач⁴. Например, сохранение углов и направлений делает проекцию Меркатора незаменимой для навигации. Однако есть несколько вариаций данной проекции: Веб–Меркатор, косая и поперечная проекции Меркатора. Первая используется в веб-приложениях ввиду того, что она хорошо подходит в качестве интерактивной карты мира, которую можно плавно масштабировать до локальных (крупномасштабных) карт. Вторая используется для картографирования территорий, которые простираются вдоль наклонной линии относительно экватора, что позволяет минимизировать искажения на больших широтах или диагонально расположенных регионах (то есть под углом к основным параллелям и меридианам). Последняя подходит для картографирования в крупном масштабе или для небольших районов, вытянутых с севера на юг. Частным примером поперечной проекции Меркатора является универсальная поперечная проекция Меркатора (UTM)(Рисунок 1). Поверхность планеты при этом делится на зоны по 6° по долготе и 8° по широте, в которых отображение на плоскость осуществляется по-разному, учитывая заданную модель эллипсоида, которая является набором параметров Земли. В пределах каждой зоны используется поперечная проекция Меркатора, которая разворачивает поверхность планеты вдоль центрального меридиана зоны, то есть внутри нее ошибка увеличивается по мере приближения к восточной и западной границам, тогда как наибольшие искажения глобально по-прежнему возникают в районе полюсов. Это позволяет минимизировать искажения формы, площади и расстояний в пределах конкретной зоны.

Применение разных моделей в зависимости от проецируемого реги-

⁴Таблица свойств проекций ESRI, URL: https://drive.google.com/file/d/1-kBEw0q6o9Fj9_bqMNyiePdIS0IBMFu0/view?usp=sharing, дата обращения: 09.05.2025

⁵Источник: https://en.wikipedia.org/wiki/Universal_Transverse_Mercator_coordinate_system#/media/File:Universal_Transverse_Mercator_zones.svg, дата обращения: 27.12.2024

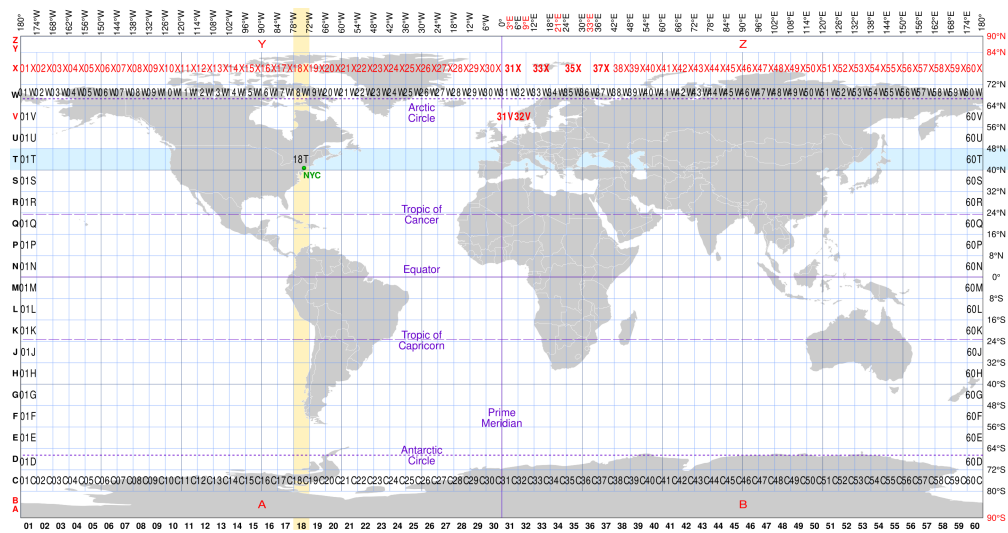


Рис. 1: Разбиение на зоны UTM⁵

она может дать более точные результаты, необходимую модель можно найти на сайте EPSG⁶ — это публичный реестр геодезических данных, систем пространственной привязки, земных эллипсоидов, преобразований координат и связанных с ними единиц измерения. У каждой модели есть свой номер и запись в формате WKT (Well-known text), представляющая векторную геометрию и описание систем координат. Например, существует много моделей для Северной Америки и отдельных регионов Европы, однако в общем случае используется универсальная для всей Земли WGS84. Стандартные зоны UTM обозначаются как 326XX для южного полушария и 327XX для северного, где последние две цифры отвечают за номер зоны по долготе.

Следует отметить, что деление на диапазоны по широтам применяется исключительно в целях описания и визуализации, тогда как при непосредственном перепроецировании координат такое разделение не играет роли и в вычислительном процессе не используется.

Учитывая вышеописанные сложности, связанные с использованием UTM-проекции в условиях покрытия сразу нескольких зон, ее применение в качестве единственной системы координат оказывается нецелесообразным. В качестве альтернативы следует рассмотреть использование единой проекции, обеспечивающей согласованность координат

⁶Публичный реестр геодезических данных, URL: <https://epsg.io/?q=>, дата обращения: 09.05.2025

по всей территории.

Координаты из разных UTM-зон также нельзя напрямую использовать совместно, так как каждая зона имеет свою систему отсчета: ось X начинается от центрального меридиана зоны.

Несмотря на то, что координаты в системе UTM формально допустимы лишь в пределах соответствующей зоны, умеренный выход за ее границы не приводит к значительным искажениям. Для оценки приемлемости такого отклонения в рамках нашей задачи далее будет проведен численный анализ возникающей ошибки.

Для оценки ошибки, возникающей при использовании координат за пределами соответствующей UTM-зоны, был проведен следующий численный эксперимент: для каждой широты от экватора до 86° с шагом в 1° вычислялось расстояние от границы до точек, находящихся на фиксированных расстояниях от границы UTM-зоны: 0.0001° , 0.001° , 0.01° , 0.1° и 0.5° . Каждая такая точка переводилась в проекцию данной зоны, несмотря на то, что географически она располагалась за ее пределами.

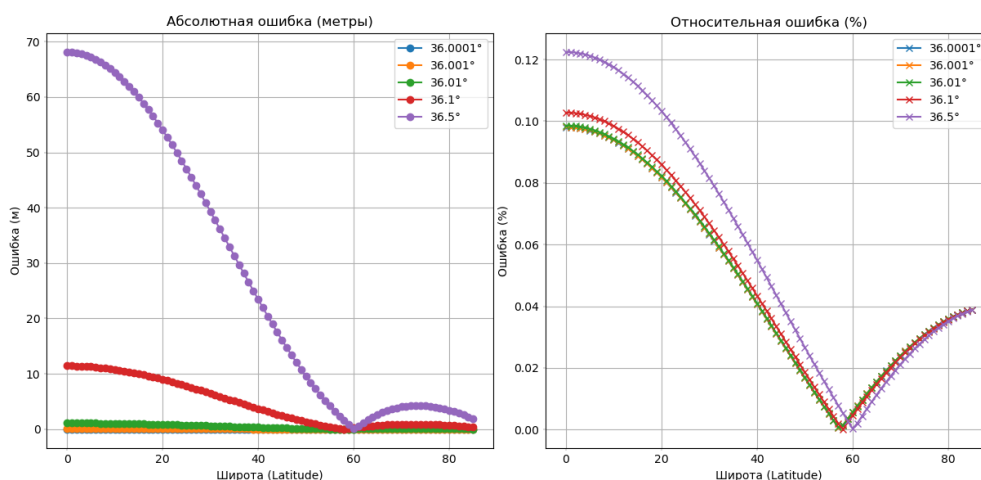


Рис. 2: Погрешность при выходе из зоны UTM⁷

Проведение данного эксперимента является достаточным, поскольку ключевым фактором, влияющим на величину искажения при проецировании, выступает широта, а не долгота. Поскольку используется эллипсоидальная модель Земли, все зоны формируются одинаково, и оценка ошибки вдоль одной долготы обеспечивает единые результаты для всех зон.

4. Реализация

4.1. Требования

1. Функциональные требования:

- (a) Обеспечение обработки входных данных, предоставляемых в формате JSON.
- (b) Обработка больших объемов географических данных в реальном времени.
- (c) Проецирование регионов, покрывающих более чем одну зону UTM.
- (d) Обеспечение совместимости перепроецированных координат.
- (e) Хранение всей истории полученных значений для анализа и уточнения.
- (f) Обновление проекции по мере накопления погрешности.

2. Нефункциональные требования

- (a) Ограничение на точность: погрешность трансформации координатных данных не должна превышать 0,05% относительной ошибки либо 20 метров абсолютного отклонения, в зависимости от контекста использования.
- (b) С практической точки зрения интерес представляет оценка погрешности на расстояниях до 100 км. между крайними точками, поскольку именно в этом диапазоне планируется применение алгоритмов. Однако для обеспечения полноты анализа и оценки устойчивости метода, необходимо также рассмотреть случаи с расстояниями до 1000 км.
- (c) Используемые библиотеки и инструменты должны распространяться под лицензиями, позволяющими их свободное использование.

- (d) Согласно требованиям заказчика разработка библиотеки должна выполняться на языке C++. Это условие объясняется высокой производительностью для обработки больших объемов геоданных в реальном времени.

В целевой системе, в рамках которой предполагается интеграция разрабатываемого решения, формат JSON используется в качестве основного средства передачи данных.

В принимаемых JSON-файлах указаны координаты четырех угловых пикселей области, выбранной в процессе геолокации. На основе этих данных определяется соответствующий участок исходного GeoTIFF-изображения, из которого извлекается UTM-координата, соответствующая центральному пикселю указанной области.

Следует отметить, что высокая точность координатных преобразований теряет практическую значимость, если превышает пространственное разрешение исходных данных. В частности, погрешность, меньшая, чем размер одного пикселя изображения, не имеет смысла как в контексте использования, так и при тестировании, так как снимки с подобной детализацией на практике недоступны.

4.2. Концепция решения

Как было сформулировано ранее, необходимо разработать решение для проецирования регионов, содержащих более чем одну зону UTM. Эту задачу теоретически можно решить тремя способами: использовать какую-либо общую проекцию для всей обзораемой зоны, научиться работать одновременно с несколькими проекциями, или, в качестве альтернативы, заранее перепроецировать исходные данные в единую систему с помощью `gdalwarp`⁸, переложив основную вычислительную нагрузку на предварительный этап обработки.

Однако перепроецирование исходного изображения на область, превышающую необходимый охват (так как система по умолчанию не име-

⁸Утилита для перепроецирования и деформации изображений, URL: <https://gdal.org/en/stable/programs/gdalwarp.html>, дата обращения: 21.05.2025

ет информации о том, какие координаты будут поступать), нецелесообразно, поскольку требует несоизмеримо больших вычислительных и временных ресурсов по сравнению с обработкой только тех координат, которые действительно участвуют в расчетах. Кроме того, такая операция приводит к ухудшению качества изображения из-за интерполяции, что, в свою очередь, снижает точность результатов и может искажать исходные данные [3].

Одновременное использование координат из двух соседних UTM-зон некорректно, поскольку каждая зона — это отдельная локальная проекция, в которой координаты считаются относительно своего центрального меридиана. Это означает, что одна и та же географическая точка будет иметь разные UTM-координаты в зависимости от выбранной зоны. При объединении координат из разных зон без приведения к общей системе теряется единая система отсчета, и, как следствие, искажаются расстояния и взаимное положение объектов.

Таким образом, необходимо выбрать альтернативный подход: вместо деления территории на зоны с различными проекциями можно задать единые правила отображения для всей интересующей области.

При подборе альтернативной картографической проекции нельзя просто опираться на признак эквидистанционности: многие проекции, формально сохраняющие расстояния, делают это лишь между отдельными точками или только вдоль определенных направлений. В то же время, отдельные равноугольные и эквивалентные проекции могут обеспечивать сохранение расстояний локально.

Например, одним из решений является применение ко всей территории проекции UTM, соответствующей центральной зоне, распространяя ее на весь участок, что позволяет минимизировать искажения в центральной части территории. Другой способ — выбрать UTM-проекцию, подходящую для той зоны, которая занимает наибольшую площадь в интересующей области.

Эти подходы выглядят перспективно, однако, как показано на рисунке 2, даже незначительный выход за пределы зоны проекции может приводить к неприемлемой погрешности в расчетах расстояний.

Так, следует использовать иную единую проекцию. Это решение обеспечивает более точное отображение для большей части региона, при этом искажения равномерны и предсказуемы в периферийных областях.

Для дальнейшего анализа была отобрана ограниченная подборка проекций, выбор которых осуществлялся на основе предварительного теоретического анализа их свойств и практической целесообразности применения в рассматриваемом контексте.

4.3. Тестирование устойчивости проекций к погрешностям

Тестирование проводилось в два этапа: на первом этапе параметры проекции подбирались индивидуально для каждой пары точек, между которыми вычислялись евклидово и геодезическое расстояния, на втором — параметры проекции фиксировались, а оценка проводилась по множеству точек, расположенных в окрестности центра проекции.

После теоретического отбора проекций остались следующие:

1. Равноугольные:

- (a) Поперечная проекция Меркатора
- (b) Стереографическая
- (c) Коническая проекция Ламберта

2. Равнопромежуточные:

- (a) Азимутальная
- (b) Коническая

На первом этапе тестирования точки перебирались в диапазоне от экватора до 76° широты. Затем, по азимутам с шагом 15° , откладывались расстояния от 1 метра до 1000 километров. В зависимости от правил параметризации проекции выбирались такие значения, чтобы либо обе точки находились на равном удалении от центра проекции,

либо, наоборот, располагались на равном расстоянии от границ области отображения.

Для решения прямой задачи использовалась библиотека GeographicLib, в которой используется метод Карнея [2], который использует разложения в ряды, что обеспечивает высокую точность даже для точек, находящихся на противоположных сторонах друг от друга.

После первого этапа тестирования были оставлены две проекции — эквидистанциальная коническая и поперечная проекция Меркатора, с которыми проводился второй этап. Первая задается минимальной и максимальной широтами охвата, вторая — центральной долготой. В рамках второго этапа при фиксированных параметрах проекции перебирались точки внутри квадрата вокруг центра проекции со стороной от 2 до 16 градусов. От этих точек, аналогично первому этапу, откладывались отрезки по азимутам с шагом 15° и длинами от 1 метра до 1000 км.

Для более точного расчета расстояния важно учитывать не только параметры эллипсоида, но и геоид, который представляет собой более точную модель формы Земли, учитывающую изменения высот рельефа. Относительно геоида ведется отсчет высот над уровнем моря. Эллипсоид используется для приближенного расчета, но не учитывает изменение высот, особенно в местах с низменностями и горными участками. Геоид же позволяет учитывать высоты, что важно для обеспечения постоянного расстояния между летающим средством и землей. Однако из-за отсутствия референсных данных нет возможности проверить эффективность данного метода.

Более того, у местностей есть дополнительный рельеф, информация о котором не находится в открытом доступе.

4.4. Моделирование

Поскольку по результатам предварительного этапа были отобраны поперечная проекция Меркатора и эквидистанциальная коническая проекция, необходимо провести моделирование работы системы с ис-

пользованием каждой из них. Цель моделирования — выявить, какая из проекций является наиболее подходящей для решаемой задачи, а также проверить работоспособность разработанной концепции в условиях, приближенных к реальной эксплуатации.

Фактически, этап моделирования представлял собой полноценную реализацию, поскольку проведение экспериментов требовало работы с той же системой, которая предполагается к использованию в дальнейшем.

Моделирование осуществляется следующим образом: точки по цепочке пачками (то есть сразу генерируется случайное количество точек, так как в реальных условиях координаты могут поступать в систему не по одной) генерируются случайно по ранее описанному принципу до тех пор, пока максимальное расстояние между ними не достигнет заданного референсного значения. Для более точного соответствия предполагаемой работе системы угол, по которому откладывается расстояние, был ограничен двумя диапазонами: сначала от 0° до 180° , а затем — от 90° до 270° . Это сделано для того, чтобы исключить чередование движений в противоположные стороны: например, сначала строго на восток, затем строго на запад. В таком случае точки могут скапливаться вблизи друг друга или возвращаться в уже пройденные области, что существенно затрудняет достижение заданного референсного значения максимального расстояния между ними. Используя результаты, полученные на предыдущем этапе, мы имеем представление о росте погрешности проецирования и можем определить момент, когда требуется обновление параметров проекции. После достижения неприемлемого уровня искажения собирается необходимая информация для построения новой проекции (средняя долгота для поперечной проекции Меркатора и максимальная и минимальная широты для конической эквидистанциальной проекции), и все ранее сгенерированные точки перепроецируются с использованием актуализированных параметров. Такой подход обеспечивает адаптивность модели к работе с протяженными территориями, включая регионы, охватывающие несколько зон UTM, при сохранении требуемой точности.

Результаты пятисот запусков представлены на Рисунке 3 и 4 в абсолютных и относительных значениях соответственно.

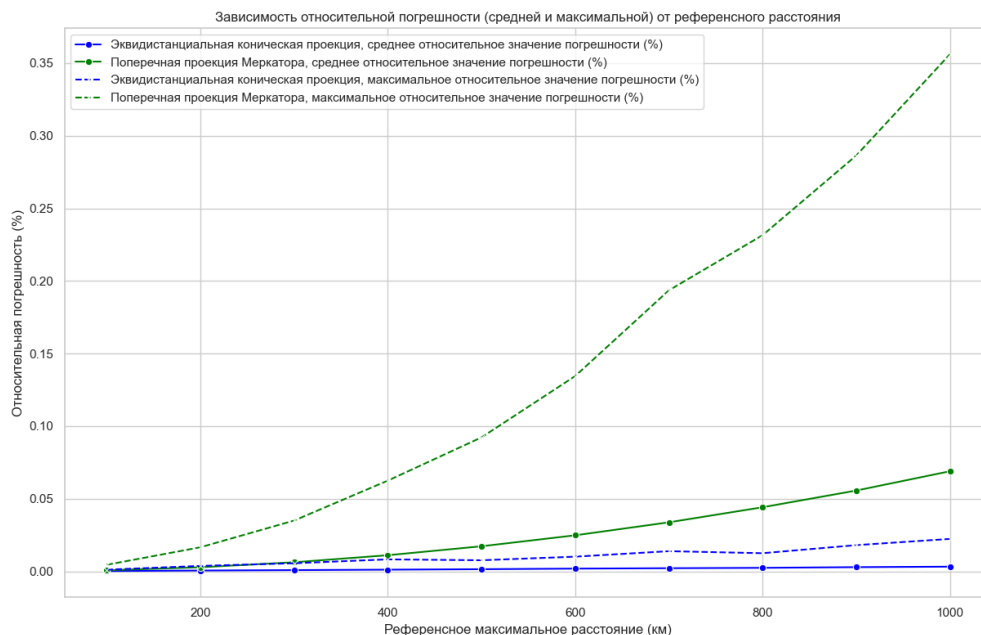


Рис. 3: Результаты моделирования при ограничении азимута от 0° до 180° в относительных значениях

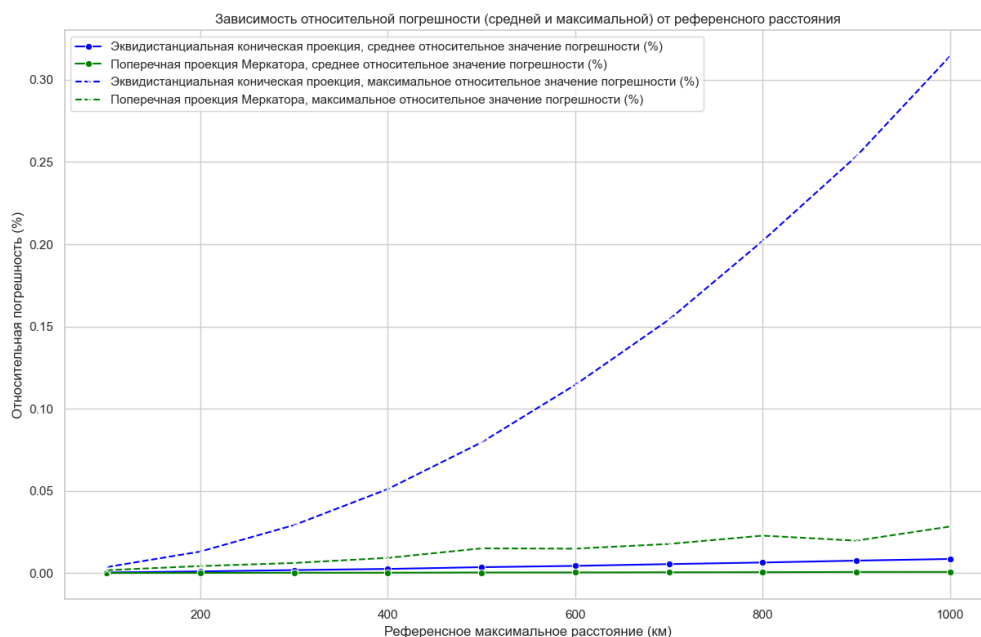


Рис. 4: Результаты моделирования при ограничении азимута от 90° до 270° в относительных значениях

Анализ графиков показывает, что эквидистанциальная коническая проекция демонстрирует значительно лучшие результаты при модели-

ровании движения, происходящего преимущественно вдоль параллелей. Это вполне объяснимо с учетом ее свойств: эквидистанциальная коническая проекция сохраняет расстояния вдоль заданных параллелей, при этом деформации распределяются симметрично относительно центральной широты, а меридианы отображаются в виде прямых, сходящихся в одной точке. Благодаря этому она обеспечивает устойчивую точность при умеренных смещениях в широтном и продольном направлениях.

Более того, при движении в направлении от 90° до 270° , ошибка в эквидистанциальной конической проекции возрастает значительно медленнее по сравнению с поперечной проекцией Меркатора, используемой для направлений от 0° до 180° .

Полученные выводы согласуются с результатами, полученными на предварительном этапе тестирования, в котором отбирались проекции на основе динамики роста искажений.

4.5. Архитектура

Для максимально точного воспроизведения работы системы были собраны соответствующие GeoTIFF-файлы из открытых источников (например, GitHub-репозитория⁹. Для подготовки данных к тестированию использовался bash-скрипт, автоматически определяющий зоны проекций, присутствующие на изображении, разделяющий исходный файл по этим зонам и выполняющий перепроецирование каждого фрагмента. Это решение выбрано из-за простоты автоматизации вызова готовых GDAL-утилит и скорости реализации. Поскольку скрипт предназначен исключительно для генерации тестовых данных, способ реализации не влияет на конечный результат. Такой подход позволяет имитировать функционирование системы с учетом специфики обработки многозональных регионов во время тестирования.

Итак, библиотека состоит из следующих компонентов.

⁹GitHub-репозитори1 с примерами GeoTIFF-файлов, URL: <https://github.com/GeoTIFF/test-data>, дата обращения: 19.05.2025

1. `PointSource` — источник точек, представляющий собой интерфейс, абстрагирующий источник данных, имеющий две реализации:
 - (a) `RandomPointGenerator` — генератор случайных точек по цепочке, когда новая точка получается при .
 - (b) `JsonReader` читает значения из JSON-файла.
2. `ProjectionManager` отвечает за выбор параметров проекции и принятие решения о ее обновлении.
3. `ProjectionEngine` — компонент, непосредственно производящий перепроецирование.
4. `DistanceAnalyzer` анализирует максимальное расстояние между точками, а также погрешности.
5. `StorageManager` обеспечивает работу с базой данных, где хранится информация об уже существующих координатах.
6. `ResultsAggregator` собирает результаты вычислений погрешностей и статистику по ним.
7. `PlotVisualiser` строит графики.
8. `SimulationRunner` инициализирует источник точек, обеспечивает их хранение, добавляет новые, вызывает `ProjectionManager` при необходимости обновления проекции, сохраняет статистику.

5. Эксперимент

5.1. Условия эксперимента

Проект должен быть протестирован и работоспособен на Ubuntu 22.04 LTS.

Учитывая, что в открытых источниках доступны изображения лишь с весьма низким пространственным разрешением (то есть с большим числом метров на пиксель), проведение эксперимента на их основе не позволило бы получить достаточно информативных или репрезентативных результатов. В связи с этим было принято решение использовать для эксперимента тот же подход, что и при моделировании — генерацию случайных точек.

5.2. Исследовательские вопросы

- Как часто нужно проводить обновление проекции?
- С какой точностью определяется значение спроецированной координаты?

5.3. Метрики

При оценке результатов работы библиотеки следует учитывать следующие метрики: точность измерений в метрах, точность измерений в процентах от референсного значения.

5.4. Результаты эксперимента

5.4.1. Частота обновления проекции

Вопрос частоты обновления параметров проекции при генерации и анализе точек имеет не только вычислительное, но и методологическое значение. На каждом этапе работы системы есть возможность перепределять параметры проекции после добавления каждой новой точки, особенно в условиях расширения обзора области. Однако на

практике такой подход приводит к избыточным вычислениям, в то время как влияние на точность остается минимальным.

Анализ результатов эксперимента показывает, что при пространственном охвате до 10 км. между крайними точками повторное перепроецирование практически не влияет на расчет новых координат. При этом погрешность трансформаций остается в пределах допустимого диапазона.

Так, оптимальной стратегией оказывается обновление параметров проекции не реже, чем один раз на каждые 10 км прироста пространственного охвата. Такой подход позволяет, с одной стороны, своевременно отследить пересечение границ зон UTM (как упоминалось ранее, критичен выход точки за пределы зоны проекции всего на 0.1 градуса по долготе, что эквивалентно приблизительно 10 км), а с другой — существенно снизить издержки на перерасчет, не жертвуя качеством метрик погрешности.

5.4.2. Точность определения значения спроецированной координаты

Расст. (км)	Ср. погр. (м)	Ср.кв. откл. (м)	Макс. погр. (м)
100	0.089	± 0.051	0.841
200	0.406	± 0.239	4.432
300	0.928	± 0.549	11.303
400	1.657	± 0.962	20.596
500	2.728	± 1.639	32.478
600	4.19	± 2.535	46.193
700	5.425	± 3.506	90.804
800	7.485	± 4.769	99.19
900	9.046	± 5.949	113.098
1000	11.098	± 7.141	149.689

Таблица 1: Погрешности определения значения спроецированной координаты

Суммарный показатель, включающий среднюю ошибку и среднеквадратическое отклонение, можно рассматривать как область, в пре-

делах которой находится истинное положение точки. Такой подход позволяет количественно оценить надежность координатной привязки, опираясь на статистические характеристики погрешности.

Помимо средних показателей, целесообразно также учитывать максимальное значение ошибки, поскольку оно позволяет оценить наихудший возможный сценарий определения координат. Это особенно важно в системах, где критична верхняя граница допустимой погрешности.

Результаты проведенного эксперимента приведены в Таблице 1.

Заключение

В ходе работы над библиотекой для работы с геоданными и изображениями в формате GeoTIFF были выполнены следующие задачи:

- Изучена проблемная область (проекции Земли):
 - Выявлены допустимые пределы использования исходной Универсальной поперечной проекции Меркатора.
 - Были проанализированы наиболее распространенные картографические проекции, изучены их математические основы, области применения, а также параметры, влияющие на характер искажения географических данных.
- Собраны требования к библиотеке.
- Созданы и реализованы алгоритмы:
 - Отобраны потенциально пригодные проекции по принципам их работы.
 - Проведены два этапа тестирований с целью оценки скорости накопления погрешности при отдалении от центра проекции.
 - Реализован алгоритм, осуществляющий автоматическое преобразование координат в новую систему при обновлении данных.
 - Была смоделирована работа системы в приближенных к реальности условиях¹⁰.
 - Был сделан вывод, что наиболее целесообразно использовать эквидистанциальную коническую проекцию.
- Спроектирована и реализована библиотека¹¹.
- Проведено итоговое тестирование библиотеки.

¹⁰GitHub-репозиторий с промежуточными экспериментами, URL: <https://github.com/chilison/zone-aware-reproj-testing-py.git>, дата обращения: 23.05.2024

¹¹GitHub-репозиторий с библиотекой, URL: <https://github.com/chilison/zone-aware-reproj-cpp.git>, дата обращения: 23.05.2024

Список литературы

- [1] Hofmann-Wellenhof Bernhard, Legat Klaus, Wieser Manfred. [Image-based navigation](#) // Navigation. — Springer Vienna, 2003. — P. 247–275. — ISBN: 9783709160787. — URL: http://dx.doi.org/10.1007/978-3-7091-6078-7_12.
- [2] Karney Charles F. F. Algorithms for geodesics // [Journal of Geodesy](#). — 2012. — . — Vol. 87, no. 1. — P. 43–55. — URL: <http://dx.doi.org/10.1007/s00190-012-0578-z> (online; accessed: 2025-05-08).
- [3] Raster Resampling for Discrete and Continuous Data. — URL: <https://gisgeography.com/raster-resampling/> (online; accessed: 2025-05-17).
- [4] Документация библиотеки GDAL для работы с геопространственными данными. — URL: https://gdal.org/_downloads/en/latest/pdf/ (online; accessed: 2024-12-23).
- [5] Документация библиотеки GeographicLib для геодезических вычислений. — URL: <https://geographiclib.sourceforge.io/C++/doc/index.html> (online; accessed: 2025-05-08).
- [6] Документация библиотеки PROJ для работы с картографическими проекциями. — URL: https://proj.org/_downloads/en/stable/pdf/ (online; accessed: 2024-12-23).
- [7] Документация библиотеки libgeotiff для работы с GeoTIFF файлами. — URL: <https://github.com/OSGeo/libgeotiff/blob/master/libgeotiff/docs/manual.tx> (online; accessed: 2024-12-23).
- [8] Документация библиотеки libtiff для работы с TIFF файлами. — URL: <https://download.osgeo.org/libtiff/doc/TIFF6.pdf> (online; accessed: 2025-05-17).