

Министерство образования и науки Российской Федерации
Санкт-Петербургский политехнический университет Петра Великого
Институт компьютерных наук и кибербезопасности
Высшая школа программной инженерии

Работа допущена к защите
Директор ВШПИ
_____ П. Д. Дробинцев
«_____» _____ 2024 г.

ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА
работа бакалавра

СРАВНЕНИЕ МЕТОДОВ СЖАТИЯ АТТРИБУТОВ ОБЛАКОВ ТОЧЕК

по направлению подготовки (специальности)

09.03.04 Программная инженерия

Направленность (профиль)

09.03.04_1 «Технология разработки и сопровождения качественного
программного продукта»

Выполнил студент гр.
5130904/00104

Поздняков А. А.

Руководитель старший
преподаватель

Фёдоров С. А.

Консультант по
нормоконтролю

Локшина Е. Г.

Санкт-Петербург
2024

САНКТ-ПЕТЕРБУРГСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ
ПЕТРА ВЕЛИКОГО

Институт компьютерных наук и технологий
Высшая школа программной инженерии

УТВЕРЖДАЮ

Директор ВШПИ

_____ П.Д. Дробинцев

”18” апреля 2024 г.

ЗАДАНИЕ

на выполнение выпускной квалификационной работы

студенту Позднякову Артемию Анатольевичу, группа 5130904/00104

1. Тема работы: Сравнение методов сжатия атрибутов облаков точек
2. Срок сдачи студентом законченной работы: 20.05.2024
3. Исходные данные по работе:
 - Документация на язык программирования Python;
 - Документация к библиотеке Open3D;
 - Документация к библиотеке NumPy;
4. Содержание работы (перечень подлежащих разработке вопросов):
 - Обоснование актуальности работы
 - Обзор существующих решений
 - Составление требований к разрабатываемому решению
 - Описание реализации
 - Анализ результатов
5. Перечень графического материала (с указанием обязательных чертежей):
6. Консультанты по работе:
7. Дата выдачи задания: 18.04.2024

Руководитель ВКР

Фёдоров С. А.

Задание принял к исполнению 18.04.2024

Студент

Поздняков А. А.

РЕФЕРАТ

Бакалаврская работа посвящена разработке подхода к оценке различных методов сжатия атрибутов облаков точек на платформе PCCArena. Дан обзор существующих алгоритмов и распространённых кодеков с открытым исходным кодом. В результате сравнительного анализа алгоритмов получены данные об эффективности (!!!) отобранных алгоритмов сжатия атрибутов облаков точек при различных данных и входных параметрах.

В рамках работы была разработана программа для вычисления качественных метрик сжатия облаков точек с использования языка Python. Предложенное решение было внедрено в существующую платформу оценки методов сжатия облаков точек PCCArena.

ABSTRACT

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus tincidunt ultrices. Lorem ipsum dolor sit amet, consectetur adipiscing elit. In hac habitasse platea dictumst. Integer tempus convallis augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet, enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus quis tortor vitae risus porta vehicula.

СОДЕРЖАНИЕ

Введение	6
Глава 1. Обзор подходов к сравнению методов сжатия облаков точек	8
1.1 Подходы к оценке качества сжатия изображений	8
1.2 Подходы к оценке качества сжатия облаков точек	14
Глава 2. Система подсчёта метрик	14
Глава 3. Результаты работы	14
Заключение	14
Список источников.....	15
Приложение А. Название приложения	18
Приложение Б. Название приложения	19

ВВЕДЕНИЕ

Облака точек. Растущая популярность (?ссылка?) технологий компьютерного зрения (CV - Computer Vision) и расширенной реальности (XR - eXtended Reality) влечёт за собой потребность в способах компактного хранения и передачи трёхмерных (далее 3Д) данных. 3Д-данные представляются в виде полигональных сеток - совокупности вершин, рёбер и граней, определяющих форму трёхмерного объекта, или облаков точек, отличающихся от последних отсутствием связей между вершинами. Дополнительно, данные о геометрической структуре объекта могут быть дополнены информацией о его внешних характеристиках.

Важной задачей является создание цифрового представления реальных объектов и сцен. 3Д-сканирование - технология, позволяющая считывать форму физического объекта и его внешних характеристик, таких как цвет или отражающая способность поверхности. В общем случае, результатом процесса 3Д-сканирования является конечное множество точек в трёхмерном пространстве[6, с. 10].

Полигональные сетки аппроксимируют непрерывную поверхность исходного объекта. В свою очередь, облака точек сохраняют мельчайшие подробности структуры поверхности объекта вплоть до миллиметра[6, с. 33]. Задача получения 3Д-модели объекта в виде полигональной сетки по имеющемуся облаку точек, считанных с его поверхности, решается методами реконструкции поверхности[6]. Для обратного преобразования достаточно удалить связи между вершинами в полигональной сетке.

Большая точность позволяет использовать облака точек для машинной обработки в системах компьютерного зрения, взаимодействующих с реальным миром. Примером подобных систем являются беспилотные автомобили, использующие лидары - локаторы, испускающие световые волны оптического

диапазона с дальнейшей регистрацией отраженных импульсов[10, с. 7]. Другим примером являются системы расширенной реальности, в данном случае облака точек используются для совмещения позиций виртуальных объектов и физических объектов, находящихся рядом с человеком[10, с. 15].

Для точного описания поверхности, облака точек должны быть достаточно плотными. Точки в облаке представляют собой дискретные образцы непрерывной поверхности, а полигональные сетки аппроксимируют данную поверхность полигонами[10, с. 4]. Всё это влечёт за собой большой размер облаков точек (может тут тоже ссылка?), в связи с чем возникает задача разработки способов компактного хранения подобных объектов.

Алгоритмы сжатия облаков точек решают задачу компактного хранения и передачи облаков точек. Популярные кодеки были реализованы в рамках проектов Draco[2] и PCL[16]. В настоящее время различными авторами предлагаются новые методы сжатия геометрической структуры облаков точек[12][15] и их атрибутов[5][14][7]. Алгоритм кодирования данных в этих решениях можно разделить на этапы реконструкции изначальной геометрической структуры объекта и кодирования атрибутов облака в соответствии с полученной структурой. Работа по стандартизации РСС-кодеков была начата MPEG в 2017 году[11], также данной группой был предложен собственный кодек и разработана тестовая модель на его основе - TMC13[3].

Большое количество постоянно появляющихся подходов к сжатию облаков точек делает актуальной задачу разработки программы для оценки работы РСС-кодеков. Подобная программа может быть использована исследователями для подсчёта метрик разрабатываемых ими кодеков.

Цель работы - разработка подхода к сравнению алгоритмов сжатия атрибутов облаков точек. В рамках данной работы необходимо решить следующие задачи:

- Проанализировать существующие РСС-кодеки
- Изучить релевантные метрики, отображающие эффективность (??) / ка-

чество (??) сжатия атрибутов облаков точек

- Разработать программу подсчёта метрик
- Получить метрики для отобранных РСС-кодеков
- Проанализировать результаты работы

ГЛАВА 1. ОБЗОР ПОДХОДОВ К СРАВНЕНИЮ МЕТОДОВ СЖАТИЯ ОБЛАКОВ ТОЧЕК

В общем случае, информация может быть сжата, если она является избыточной. Сжатие без потерь основано на уменьшения избыточности информации. В сжатии с потерями вводится новое понятие - нерелевантная информация, то есть такая информация, которая слабо влияет на восприятие изображения человеком. Сжатие с потерями, таким образом, основано не только на уменьшении избыточности, но и на определении нерелевантной информации и уменьшении количества подобной информации[1, с. 265].

В случае, когда визуальная информация предназначена прежде всего для восприятия человеком, единственной корректной метрикой визуального качества изображения является субъективная оценка некоторой отобранной группой людей[9]. Данный подход, однако, не является машинным и, следовательно, не подлежит автоматизации. Для решения данной проблемы были разработаны метрики, помогающие автоматически предсказать *воспринимаемое качество* (*perceived quality*) изображения. К подобным метрикам относятся метрики качества изображения SSIM и VIF(Visual Image Fidelity).

1.1. Подходы к оценке качества сжатия изображений

Потребность в оценке качества сжатия визуальной информации возникла после изобретения методов сжатия подобной информации. Многие из качественных метрик сжатия 3Д-данных, были изначально разработаны для оценки качества сжатия обычных изображений.

Введем понятие реконструированного изображения. Пусть оригинальное изображение A было закодировано в формате F_1 . Закодируем оригинальное

изображение с использованием алгоритма сжатия и получим сжатое изображение B , закодированное в формате F_2 . Теперь произведем декомпрессию изображения B и закодируем его с помощью формата F_1 . Полученное изображение C в формате F_1 - реконструированное изображение.

Метрики качества для медиа могут быть классифицированы согласно тому, с чем сравнивается реконструированное изображение[9]:

- С оригинальным, неискаженным изображением
- С набором признаков, извлеченных из оригинального изображения
- С самим собой (!!!)

Другая классификация предполагает разделение метрик по признаку того, на чём они основаны[13, с. 6]:

- Математически-обоснованные метрики, учитывающие лишь интенсивность искажения. К подобным метрикам относятся среднеквадратичная ошибка (MSE) и пиковое отношение сигнал-шум (PSNR)
- Низкоуровневые метрики, которые учитывают видимость искажений, как, например функции чувствительности контраста (CSF)
- Высокоуровневые метрики, которые основаны на гипотезе о том, что человеческое зрение адаптировано к извлечению структурной информации из изображения. К подобным метрикам относится индекс структурной схожести (SSIM) и визуальная точность изображения (VIF)

Математически-обоснованные метрики качества

Одной из стандартных метрик качества сжатия изображений является пиковое отношение сигнал-шум (Peak Signal to Noise Ratio - далее PSNR). Данная метрика достаточно проста для вычисления, однако она имеет лишь ограниченное, приблизительное отношение к ошибкам, которые воспринимает человеческий глаз. Иными словами, большее значение PSNR означает меньшее расхождение между оригинальным и сжатым изображением, но не гарантирует, что

зрителям понравится реконструированное изображение[1, с. 279].

Обозначим пиксели исходного изображения как P_i , а пиксели реконструированного изображения как Q_i (где $0 \leq i \leq n$). Для начала определим понятие среднеквадратичной ошибки (Mean Squared Error - далее MSE) между двумя изображениями как

$$MSE = \frac{1}{n} \sum_{i=1}^n (P_i - Q_i)^2 \quad (1.1)$$

Иными словами, среднеквадратичная ошибка для изображений является суммой квадратов ошибок для каждого из пикселей. Тогда метрика PSNR может быть определена как

$$PSNR = 10 \log_{10} \frac{\max_i |P_i|^2}{MSE} \quad (1.2)$$

При этом $\max_i |P_i|$ - пиковое значение сигнала. Для чёрно-белого изображения пиковым значением является 1, для изображения в оттенках серого при глубине 8 бит на пиксель данное значение равно 255. Так как используется логарифм отношения, результирующее значение измеряется в децибелах.

Метрика PSNR имеет значение лишь в сравнении с показателями данной метрики для того же (или схожего) кодека и тех же входных данных при отличных входных параметрах[8]. Несмотря на это, данная метрика продолжает использоваться исследователями для сравнения качества работы существующих кодеков с предлагаемыми ими решениями[4].

Другой похожей метрикой является отношение сигнала к шуму (Signal to Noise Ratio - далее SNR). В данном случае, рассматривается не пиковое значение сигнала, а среднеквадратичное.

$$SNR = 10 \log_{10} \frac{\frac{1}{n} \sum_{i=1}^n P_i^2}{MSE} \quad (1.3)$$

Значение отношения сигнала к шуму квантования (Signal to Quantization

Noise Ratio - далее SQNR) представляет собой меру влияния квантования на качество сигнала. Данную метрику можно определить как отношение мощности сигнала к разнице между сигналом после квантования и оригинальным значением сигнала.

$$SQNR = 10 \log_{10} \frac{\text{signal power}}{\text{quantization error}} \quad (1.4)$$

Высокоуровневые метрики качества

Математически-обоснованные метрики основаны на предположении, что уменьшение воспринимаемого качества изображения напрямую связано с величиной шума. Так, например, MSE даёт объективную оценку мощности (?) шума, однако два зашумленных изображения с одинаковым значением MSE могут ошибки разного рода, некоторые из которых являются более заметными чем другие. Для решения данной проблемы были предложены высокоуровневые метрики качества изображений.

Здесь можно стырить иллюстрацию с лодками из[9].

Рассмотрим индекс структурной похожести (SSIM), данная метрика была предложена в качестве способа предсказания воспринимаемого качества изображения. Она представляет собой более точную и соответствующую человеческому восприятию метрику для оценки качества сжатия изображений, поскольку учитывает структурные и текстурные аспекты изображения, в то время как MSE и PSNR ориентированы на разницу в значениях пикселей без учета визуальных особенностей.

SSIM (Structural Similarity Index)

вычисляется сравнением трех основных аспектов изображений: яркости (luminance), контрастности (contrast) и структуры (structure). Вот формула для вычисления SSIM:

$$\text{SSIM}(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (1.5)$$

где:

- x и y - сравниваемые изображения
- μ_x и μ_y - средние значения пикселей изображений x и y
- σ_x^2 и σ_y^2 - дисперсии значений пикселей изображений x и y
- σ_{xy} - ковариация между значениями пикселей изображений x и y
- c_1 и c_2 - константы для обеспечения устойчивости деления

Этот индекс обычно принимает значения от -1 до 1 , где 1 указывает на идеальное сходство между изображениями, а значения ближе к -1 указывают на более сильные различия.

SSIM сравнивает локальные окрестности пикселей изображений, а не их абсолютные значения. Он учитывает не только яркость и контрастность, но и структуру изображения, что делает его более подходящим для оценки качества изображений с точки зрения человеческого восприятия

VIF (Visual Information Fidelity)

. VIF обеспечивает более точную оценку качества сжатия, так как учитывает не только структурные аспекты изображения, но и сложные визуальные особенности, такие как текстуры, края и детали. Это делает VIF более подходящей метрикой для оценки реального восприятия качества изображения человеком.

Вычисление Visual Information Fidelity (VIF) включает несколько этапов, включающих оценку сходства между двумя изображениями с учетом их визуальной информации. Вот общий алгоритм вычисления VIF:

- Разбиение изображений на блоки: Сначала изображения разбиваются на небольшие блоки пикселей. Обычно используются квадратные блоки определенного размера.
- Вычисление локальных статистических параметров: Для каждого блока

изображения вычисляются локальные статистические параметры, такие как среднее значение, дисперсия и ковариация пикселей.

- Вычисление локальных масштабных параметров: Для каждого блока изображения также вычисляются локальные масштабные параметры, которые оценивают структурную информацию в блоке.
- Вычисление глобальных статистических параметров: Глобальные статистические параметры вычисляются на основе суммирования или усреднения локальных параметров по всем блокам изображений.
- Вычисление VIF: Затем производится расчет VIF, используя локальные и глобальные параметры. В общем, VIF представляет собой взвешенную сумму сходства между локальными статистическими параметрами двух изображений.
- Нормализация VIF: Иногда VIF может быть нормализован для получения значения в диапазоне от 0 до 1, где 1 указывает на идеальное сходство между изображениями.

Это общий алгоритм, и реальная реализация VIF может включать дополнительные детали и оптимизации. Однако основная идея заключается в том, чтобы учитывать как локальные, так и глобальные статистические параметры изображений для оценки их визуального сходства.

В целом, VIF обеспечивает более глубокую и точную оценку качества сжатия изображений, учитывая разнообразные визуальные особенности и взаимосвязи между блоками изображения, что делает его ценным инструментом при сравнении различных методов сжатия изображений.

Netflix VMAF

1.2. Подходы к оценке качества сжатия облаков точек

mpeg-pcc-dmetric

GeoCNNv1 geo_dist

PCCArena

ГЛАВА 2. СИСТЕМА ПОДСЧЁТА МЕТРИК

ГЛАВА 3. РЕЗУЛЬТАТЫ РАБОТЫ

ЗАКЛЮЧЕНИЕ

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

СПИСОК ИСТОЧНИКОВ

1. *(auth.) P. D. S.* Data Compression: The Complete Reference. — Springer, 2007. — ISBN 9781846286025; 1846286026; 9781846286032; 1846286034. — URL: [libgen . li / file . php ? md5 = d29712c14c1f0a628fa5c85329bbafb0](http://libgen.li/file.php?md5=d29712c14c1f0a628fa5c85329bbafb0).
2. Draco [Электронный ресурс]. — URL: <https://github.com/google/draco> (дата обр. 08.05.2024).
3. TMC13 [Электронный ресурс]. — URL: <https://github.com/MPEGGroup/mpeg-rcc-tmc13> (дата обр. 08.05.2024).
4. Toward A Practical Perceptual Video Quality Metric [Электронный ресурс]. — URL: <https://netflixtechblog.com/toward-a-practical-perceptual-video-quality-metric-653f208b9652> (дата обр. 10.05.2024).
5. Attribute compression of 3D point clouds using Laplacian sparsity optimized graph transform / Y. Shao [и др.] // 2017 IEEE Visual Communications and Image Processing (VCIP). — IEEE, 12.2017. — DOI: 10.1109/vcip.2017.8305131. — URL: <http://dx.doi.org/10.1109/VCIP.2017.8305131>.
6. *Bellocchio F.* 3D Surface Reconstruction: Multi-Scale Hierarchical Approaches. — 1-е изд. — Springer, 2013. — ISBN 9781461456315; 1461456312; 9781461456322; 1461456320. — URL: [libgen . li / file . php?md5=4c633bfaca5acbacd297030a9860c540](http://libgen.li/file.php?md5=4c633bfaca5acbacd297030a9860c540).
7. Cluster-based two-branch framework for point cloud attribute compression / L. Sun [и др.] // The Visual Computer. — 2023. — Нояб. — ISSN 1432-2315. — DOI: 10.1007/s00371-023-03146-9. — URL: <http://dx.doi.org/10.1007/s00371-023-03146-9>.

8. *Huynh-Thu Q., Ghanbari M.* Scope of validity of PSNR in image/video quality assessment // *Electronics Letters*. — 2008. — T. 44, № 13. — С. 800. — ISSN 0013-5194. — DOI: 10.1049/el:20080522. — URL: <http://dx.doi.org/10.1049/el:20080522>.
9. Image quality assessment: from error visibility to structural similarity / Z. Wang [и др.] // *IEEE Transactions on Image Processing*. — 2004. — T. 13, № 4. — С. 600—612. — DOI: 10.1109/TIP.2003.819861.
10. *Kuo(auth.) S. L. Z. K.-C. J.* 3D Point Cloud Analysis: Traditional, Deep Learning, and Explainable Machine Learning Methods. — 1st ed. — Springer, 2021. — ISBN 9783030891794; 3030891798; 9783030891800; 3030891801. — URL: libgen.li/file.php?md5=36c343c73075c1e6e388f26d48e36be2.
11. *MPEG.* Call for Proposals for Point Cloud Compression V2 : тех. отч. / ISO/IEC. — Hobart, AU, 04.2017. — MPEG2017/N16763.
12. Multiscale Point Cloud Geometry Compression / J. Wang [и др.]. — 2020. — DOI: 10.48550/ARXIV.2011.03799. — URL: <https://arxiv.org/abs/2011.03799>.
13. *Pedersen M., Hardeberg J. Y.* — 2012. — DOI: 10.1561/06000000037.
14. Point Cloud Attribute Compression via Successive Subspace Graph Transform / Y. Chen [и др.] // 2020 IEEE International Conference on Visual Communications and Image Processing (VCIP). — IEEE, 12.2020. — DOI: 10.1109/vcip49819.2020.9301784. — URL: <http://dx.doi.org/10.1109/VCIP49819.2020.9301784>.
15. *Quach M., Valenzise G., Dufaux F.* Improved Deep Point Cloud Geometry Compression. — 2020. — arXiv: 2006.09043 [cs.CV].

16. *Rusu R. B., Cousins S.* 3D is here: Point Cloud Library (PCL) // IEEE International Conference on Robotics and Automation (ICRA). — Shanghai, China : IEEE, 05.2011.

ПРИЛОЖЕНИЕ А. НАЗВАНИЕ ПРИЛОЖЕНИЯ

Текст приложения 1

ПРИЛОЖЕНИЕ Б. НАЗВАНИЕ ПРИЛОЖЕНИЯ