



ИПМ им.М.В.Келдыша РАН • Электронная библиотека

Препринты ИПМ • Препринт № 79 за 2022 г.



ISSN 2071-2898 (Print)
ISSN 2071-2901 (Online)

**М.С. Железов, Е.В. Ляпустин,
А.В. Москаленко, Д.С. Ватолин,
А.Г. Волобой**

**Выбор лучшей проекции для
видео 360°**

Статья доступна по лицензии
Creative Commons Attribution 4.0 International



Рекомендуемая форма библиографической ссылки: Выбор лучшей проекции для видео 360° / М.С. Железов [и др.] // Препринты ИПМ им. М.В.Келдыша. 2022. № 79. 9 с.
<https://doi.org/10.20948/prepr-2022-79>
<https://library.keldysh.ru/preprint.asp?id=2022-79>

**Ордена Ленина
ИНСТИТУТ ПРИКЛАДНОЙ МАТЕМАТИКИ
имени М.В.Келдыша
Российской академии наук**

**М.С. Железов, Е.В. Ляпустин, А.В. Москаленко,
Д.С. Ватолин, А.Г. Волобой**

Выбор лучшей проекции для видео 360°

Москва — 2022

Железов М.С., Ляпустин Е.В., Москаленко А.В., Ватолин Д.С., Волобой А.Г.

Выбор лучшей проекции для видео 360°

В статье сравнивается эффективность различных проекций для различных 360-градусных видео и кодеков, включая AV1, HEVC, VP9 и AV. Контент в видео 360° является особенно сложным, поскольку использование кодеков при его репроекции увеличивает количество ошибок видеопотока. По этой причине были протестированы разные проекции, чтобы определить их стабильность при разных способах сжатия. Результаты были получены посредством измерений качества на основе таких показателей, как PSNR, SSIM и VMAF, а также субъективной оценки статической области просмотра.

Ключевые слова: сферическое видео, видео 360, кодирование видео, AVC, HEVC, AV1, равнопромежуточная проекция, кубическая карта, бочкообразная проекция

Maksim Sergeevich Zhelezov, Eugene Vladislavovich Lyapustin, Andrey Viktorovich Moskalenko, Dmitriy Sergeevich Vatolin, Alexey Gennadievich Voloboy

Choosing the best projection for 360 video

This paper compares the effectiveness of different projections for different 360-degree videos and codecs, including AV1, HEVC, VP9, and AV. Content for 360 video is especially challenging because using codecs on reprojections magnifies video-stream errors. For this reason, Different projections have been tested to determine their stability over different encodings. Results came from performance measurements using end-to-end metrics such as PSNR, SSIM and VMAF as well as subjective assessment of static viewports.

Key words: spherical video, 360 video, video coding, AVC, HEVC, AV1, ERP, cubemap, barrel

Исследование выполнено за счет гранта Российского научного фонда № 22-21-00478, <https://rscf.ru/project/22-21-00478/>

Введение

В последние годы вырос интерес к видео 360°. Этот формат позволяет зрителям свободно просматривать различные части сцены. Поскольку сохранять видео в сферическом формате чрезвычайно сложно, их необходимо отображать на плоскость, однако не существует идеальной и универсальной проекции сферы. Все подобные методы преобразования вносят искажения, но позволяют использовать те же инструменты сжатия, что и при обработке обычного видео. Хотя кодеки настроены на работу с пространственно непрерывным видео, проекция сферы в эту категорию не попадает. В нашей работе мы изучаем, как кодеки сжимают 360-градусное видео в разных проекциях, и определяем лучшую проекцию для обработки этого контента.

Обзор литературы

Как уже было отмечено выше, сфера может быть преобразована в плоскость с помощью различных типов проекций. Многие из этих методов отражены в работе Совместной группы по исследованию видео (JVET) [1], здесь же мы кратко рассмотрим некоторые из них.

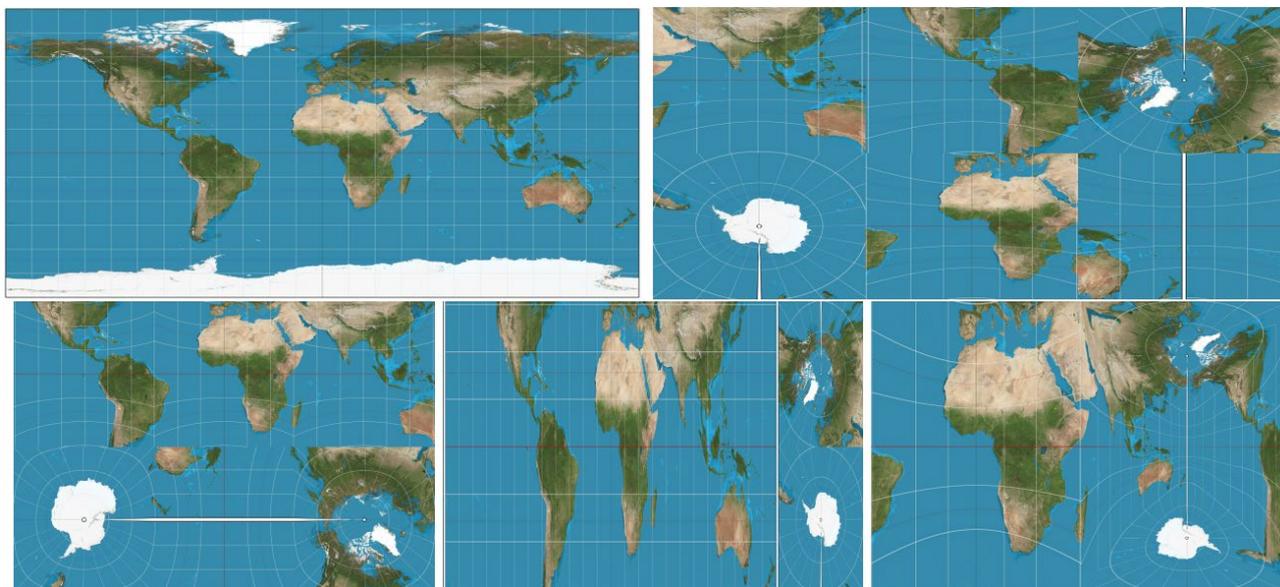


Рис. 1. Равнопромежуточная проекция Земли, отображённая в различные проекции. Верхний ряд слева направо: равнопромежуточная проекция, кубическая карта. Нижний ряд слева направо: равноугольная кубическая карта, бочкообразная проекция, проекция усечённой квадратной пирамиды

Наиболее распространённой является так называемая равнопромежуточная проекция (ERP). Она отображает широту и долготу в равной степени на двумерную плоскость. Результат оказывается сильно растянут на северном и

южном полюсах, поэтому точки в этих областях кодируются большим числом пикселей, чем экватор, который наиболее важен для зрителей.

Другим способом отображения является кубическая карта (CMP). Этот подход проецирует сферу на описанный куб, разворачивая его грани на двумерной плоскости порядке 3×2 .

Равноугольная кубическая карта (EAC) [2], предложенная компанией Google и используемая на платформе YouTube, представляет собой модифицированную форму кубической карты. Эта проекция использует равные углы для проецирования граней куба на плоскость, поэтому количество пикселей на краях граней пропорционально количеству пикселей в центре. Кроме того, расположение граней фиксируется таким образом, чтобы свести к минимуму разрывы между ними. EAC добавляет отступы к краям граней для интерполяции пикселей при создании областей просмотра.

Бочкообразная проекция [3] добавляет середину равнопромежуточной проекции к северному и южному полюсам кубической карты, устраняя этим растяжение полюсов, которое создаёт ERP.

Проекция усечённой квадратной пирамиды (TSP) выделяет центральную область просмотра, потому что многие люди смотрят на вид спереди. Здесь используется пирамидальная компоновка, в которой области разных сторон неравны между собой и больше места отводится виду спереди.

Рисунок 1 иллюстрирует описанные виды проекций.

Toriwala и др. [4] сравнили различные кодеки и проекции для видео 360° . Они использовали схему, предложенную JVET, и протестировали 4 метода сжатия в равнопромежуточной проекции и проекции вращающейся сферы. Результаты показывают, что кодек libaom уступает HEVC HM 16.15, JVET JEM 6 и x265. Однако они не сравнивают другие проекции. В статье Уен и др. [5] предлагается новый метод субъективного исследования, который применяется к различным проекциям, в том числе описанным выше.

Основным вкладом настоящей работы является представление крупнейшего сравнения кодеков и проекций для видео 360° в различных сценариях. В дополнение к объективным результатам мы провели также крупномасштабное субъективное сравнение с новой методологией выбора областей просмотра для субъективной оценки.

Экспериментальная оценка

В экспериментах мы использовали десять видеопоследовательностей с разрешением 4K длительностью 300 кадров из набора данных Vimeo. В видеороликах использовались как стационарные, так и динамические камеры, а также присутствовали движущиеся объекты. На рисунке 2 показано распределение метрики SI-TI для видео в различных проекциях. Мы можем наблюдать, что разница между проекциями существует, однако она не так велика. Мы выбрали наиболее распространённые проекции и протестировали

их под другими проекциями. Для их преобразования был использован фильтр FFMPEG v360.

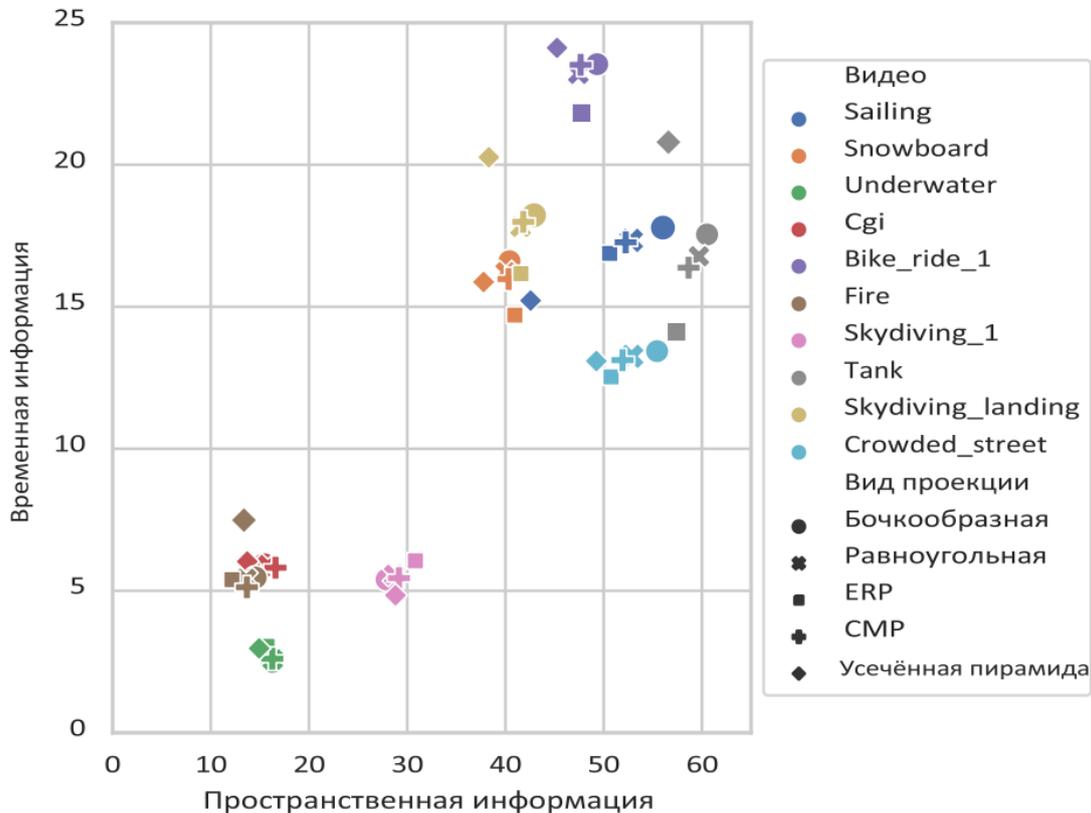


Рис. 2. Распределение метрики Si-Ti

В нашей работе были рассмотрены 5 кодировщиков: WebM Project VP9 версии 1.10 (VP9), AOMedia Project AV1 версии 3.1.1 (libaom), rav1e версии 0.5.0-alpha (rav1e), x264 версии 0.164.x (x264) и HEVC версии 3.5 (x265), — которые входные видео с битрейтом 2, 4, 6, 10, 14, 18 и 20 Мбит/с.

Мы настроили кодеки так, чтобы они соответствовали скорости кодирования 1 кадр/с на компьютере с процессором Intel Core i5-2320 и 16 ГБ ОЗУ.

Мы использовали среду для проведения тестов, рекомендованную JVET [6]. Но, поскольку в различных библиотеках оценки объективного качества отсутствует поддержка для усечённой квадратной пирамиды и бочкообразной проекции, мы не смогли измерить кросс-форматные и сквозные показатели. Однако мы измерили качество в динамической области просмотра, используя объективные и субъективные метрики. В ходе эксперимента исходное видео было отображено и уменьшено в равнопромежуточной проекции с сохранением 75% исходного разрешения. Этот процесс уменьшает влияние исходной проекции. Следующим нашим шагом было сжатие видео посредством кодировщиков, а затем применение метрик к кодекам и областям просмотра. В результате получается две области просмотра. Первая идёт слева направо

через центр, а вторая снизу вверх через вид сзади. Наше сравнение кодеков и областей просмотра измеряло показатели PSNR, SSIM и VMAF.

Для нашей субъективной оценки области просмотра мы выбрали онлайн-платформу Subjectify.us. Это позволило нам собрать более 22 000 ответов от более чем 1 500 участников. Каждому участнику показывали 15 пар расположенных рядом центральных областей просмотра и просили выбрать лучшую из каждой пары или указать, что они примерно равны. Чтобы сократить стоимость сравнения, мы использовали только 3 видеозаписи: Cgi, Bike_ride_1 и Underwater, — которые представляют различные типы контента, такие как компьютерная графика, высокодинамичные и статические камеры. Однако мы избегали тестирования этих роликов на полном диапазоне битрейтов, используя вместо этого уменьшенный диапазон: 4, 10 и 18 Мбит/с. После этого алгоритм Брэдли–Терри [7] преобразовал данные для получения субъективных оценок.

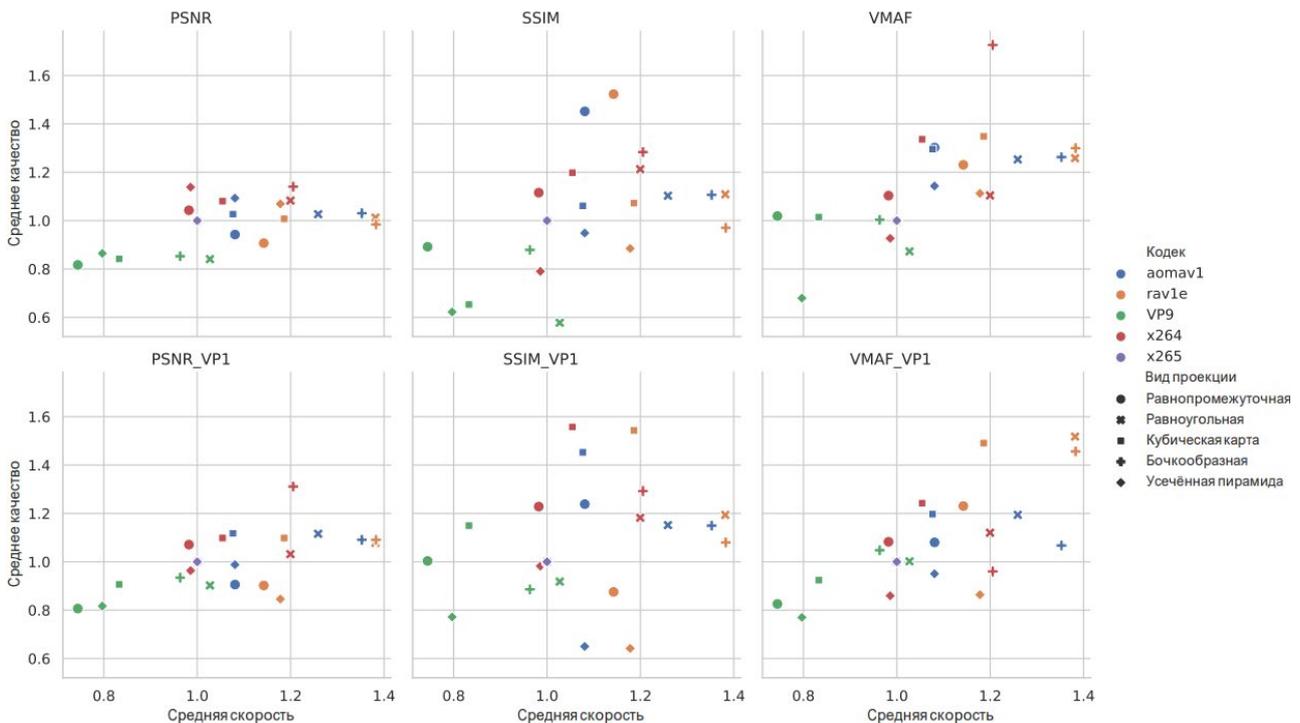


Рис. 3. BSQ-rate

Чтобы представить наши результаты, мы используем BSQ-rate [8]. Рисунок 3 сравнивает различные кодеки и проекции. Среднее качество означает, что результат лучше или хуже, чем у других записей по отношению к тому же битрейту. Чем оно выше, тем больше битрейта сохраняется. То же самое со средней скоростью, которая показывает, насколько быстрее или медленнее метод по сравнению с другими. Мы видим, что rav1e и x264 опережают по качеству другие кодеки. Проекция сильно различается в зависимости от метрики и кодека, однако высокие результаты показывают равнопромежуточная, бочкообразная проекция и кубическая карта, при этом

проекция усечённой квадратной пирамиды показывает худший результат из всех.

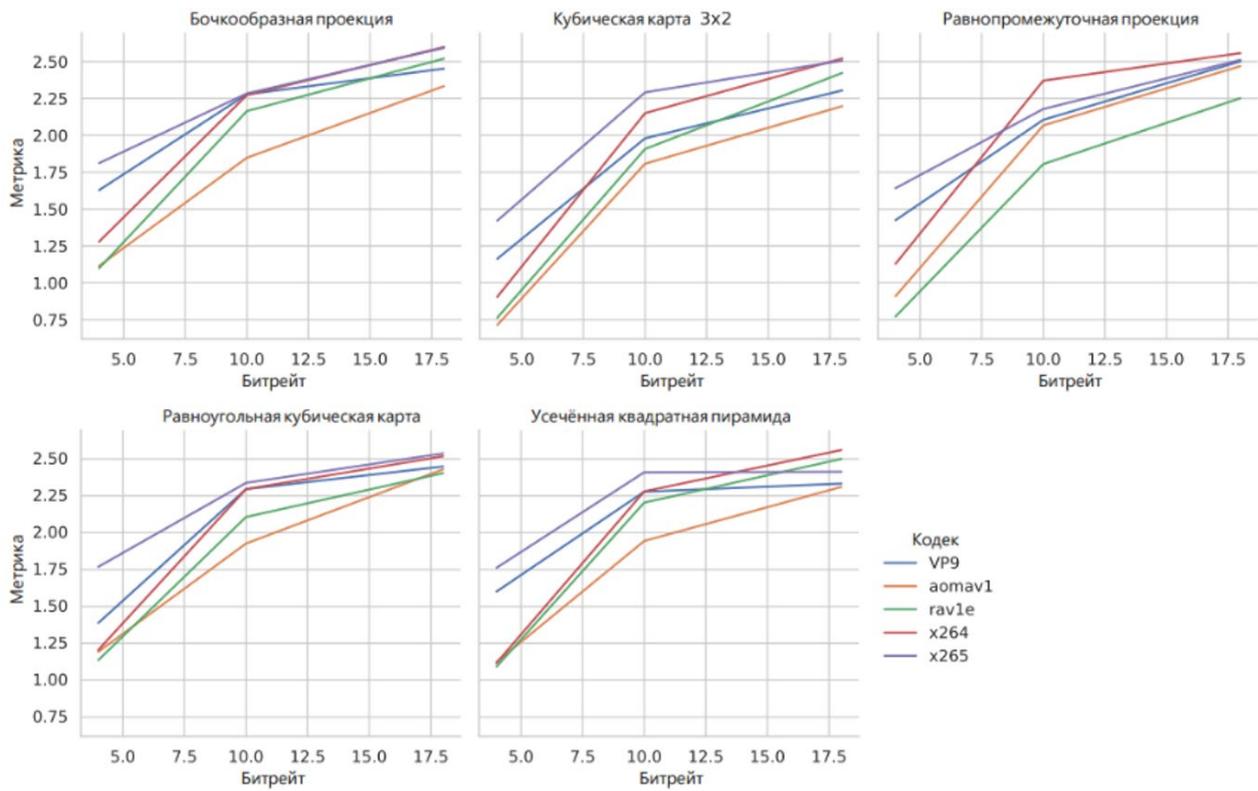


Рис. 4. RD-кривые субъективных результатов для разных проекций

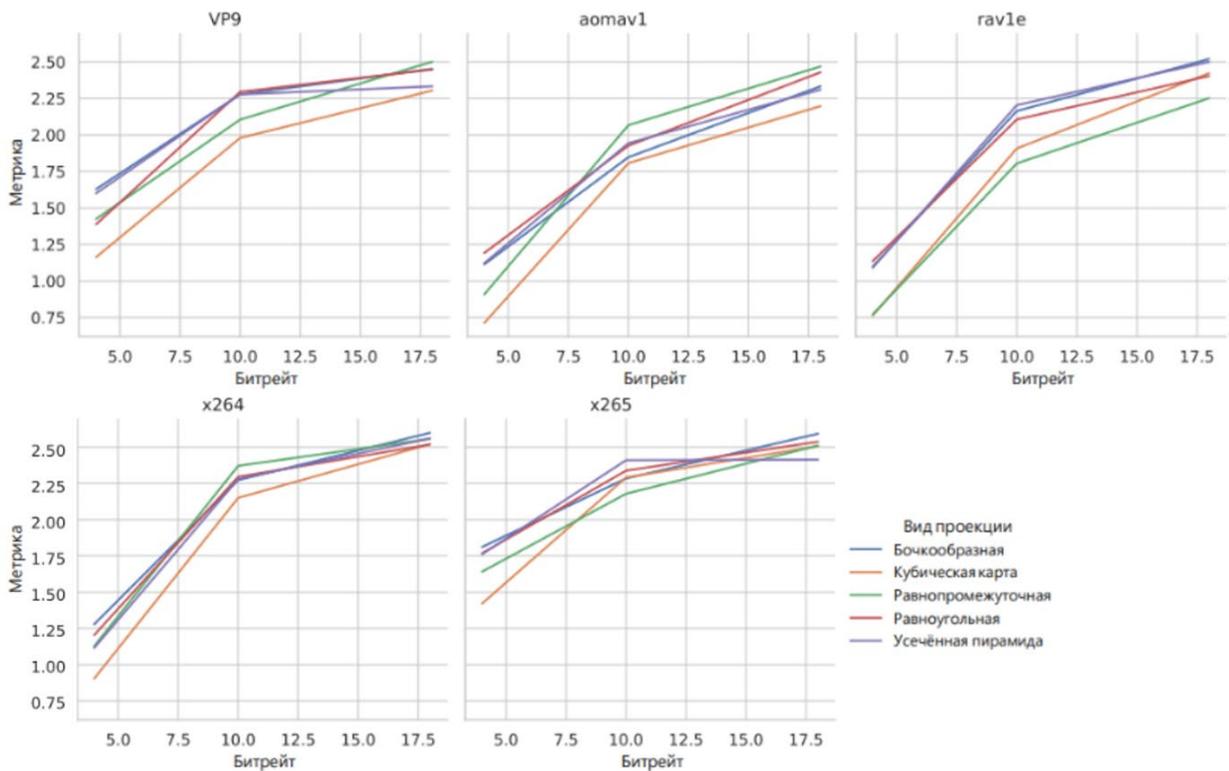


Рис. 5. RD-кривые субъективных результатов для разных кодеков

Для субъективной оценки мы анализируем RD-кривые. На рисунке 4 показано, что люди предпочли кодек x265 почти для всех проекций, а aomav1 был последним выбором среди кандидатов. Анализ проекций на субъективных данных, представленных на рисунке 5, показывает, что кубическая карта, а также равнопромежуточная и бочкообразная проекция практически неразличимы. При этом проекция усечённой квадратной пирамиды показывает стабильно плохие результаты практически с любым представленным кодеком.

Заключение

Мы обнаружили, что 360-градусное видео является сложным приложением для видеокодеков, поскольку оно отличается от их привычной области использования. Однако всё ещё необходимо кодирование высококачественных кадров, в частности в локальных областях, таких как области просмотра и полюса. В нашем анализе новые кодеки rav1e и x264 превзошли другие алгоритмы в тестировании постоянного качества, однако это также показывает нам, что они на самом деле не полностью оптимизированы для видео 360°. Если говорить о проекциях, то почти все они показали себя хорошо, за исключением усечённой квадратной пирамиды, которая показывает плохие результаты во всех сравнениях. Перед дальнейшим тестированием необходимо подтвердить эти выводы, но наши первоначальные результаты согласуются с другими сравнениями и нашим субъективным анализом.

Библиографический список

- [1] Yan Ye Boyce Jill. JVET-J1010: JVET common test conditions and software reference configurations, 07, 2018.
- [2] Chip Brown. Bringing pixels front and center in VR video, 2017.
- [3] Renbin Peng Shannon Chen, Evgeny Kuzyakov. Enhancing high-resolution 360 streaming with view prediction, 2017.
- [4] Pankaj Topiwala, Wei Dai, Madhu Krishnan, Adeel Abbas, Sandeep Doshi, and David Newman. Performance comparison of av1, hevc, and jvet video codecs on 360 (spherical) video / in: Applications of Digital Image Processing XL. International Society for Optics and Photonics, 2017, vol. 10396, p. 1039609. <https://doi.org/10.1117/12.2272119>
- [5] Tran Thi Hai Uyen, Oh-Jin Kwon, Seungcheol Choi, and Ikram Hussain. Subjective assessment of 360° image projection formats // IEEE Access, vol. 8, pp. 33588–33599, 2020. <https://doi.org/10.1109/access.2020.2974010>
- [6] Boyce Jill, Karsten Suehring, Xiang Li, and Vadim Seregin. JVET-R0006: JVET algorithm descriptions of projection format conversion and video quality metrics in 360lib version 7, 07, 2018.
- [7] Ralph Allan Bradley and Milton E Terry. Rank analysis of incomplete block designs: I. the method of paired comparisons // Biometrika, vol. 39, no. 3/4, pp. 324–345, 1952. <https://doi.org/10.1093/biomet/39.3-4.324>

[8] Anastasia V. Zvezdakova, Dmitriy L. Kulikov, Sergey V. Zvezdakov, and Dmitriy S. Vatolin. Bsqr-rate: a new approach for video-codec performance comparison and drawbacks of current solutions // Programming and computer software, vol. 46, no. 3, pp. 183–194, 2020.
<https://doi.org/10.1134/s0361768820030111>

Оглавление

Введение	3
Обзор литературы.....	3
Экспериментальная оценка	4
Заключение.....	8
Библиографический список.....	8