

Как улучшить сжатие на 20% с помощью машинного обучения, не тратя 300+ машино-лет расчетов

*Дмитрий Ватолин
Максим Смирнов*

Effective Video Transcoding
consulting@evt.guru

- Компания Effective Video Transcoding — один из стартапов членов видеогруппы лаборатории
- Мы занимаемся оптимизацией видеокодеков 21 год, сжатием 25 лет, оценкой качества кодеков 19 лет



Исследования по качеству, обработке и сжатию видео

Сравнения видеокодеков

- Объективные сравнения (SSIM, VMAF, PSNR и другие метрики)
- Субъективные сравнения
- Сравнения облачных сервисов транскодирования и др.

Разработка и анализ современных метрик качества видео

- Video Quality Measurement Tool и ПО Perceptual Quality
- Бенчмарк метрик качества
<https://videoprocessing.ai/benchmarks/video-quality-metrics.html>
- Анализ устойчивости метрик к атакам: устойчивость VMAF и других метрик
- Разработка метрик: video conferencing, blocking и blur

Технологии для улучшения сжатия и повышения качества видео

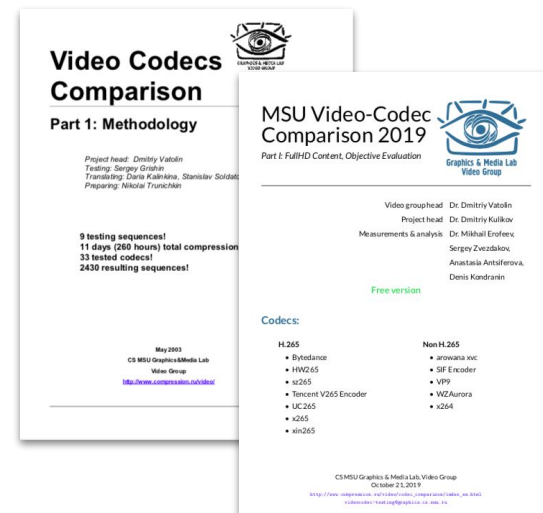
- Тюнинг кодеков, анализ сильных и слабых сторон
<https://evt.guru/>
- SuperResolution для восстановления деталей: методы, метрики, бенчмарк
[SuperResolution+compression benchmark](https://superresolution.com/compression-benchmark)
- Сжатие с использованием карт салиентности
<http://compression.ru/video/savam/>
- Видео маттинг <https://videomattng.com/>
- Восстановление видео и инпейнтинг
- Деинтерлейсинг
- Денойзинг
- Сжатие карт глубины
- Детектирование смены сцен: новый метод и бенчмарк

MSU Video Codecs Comparisons

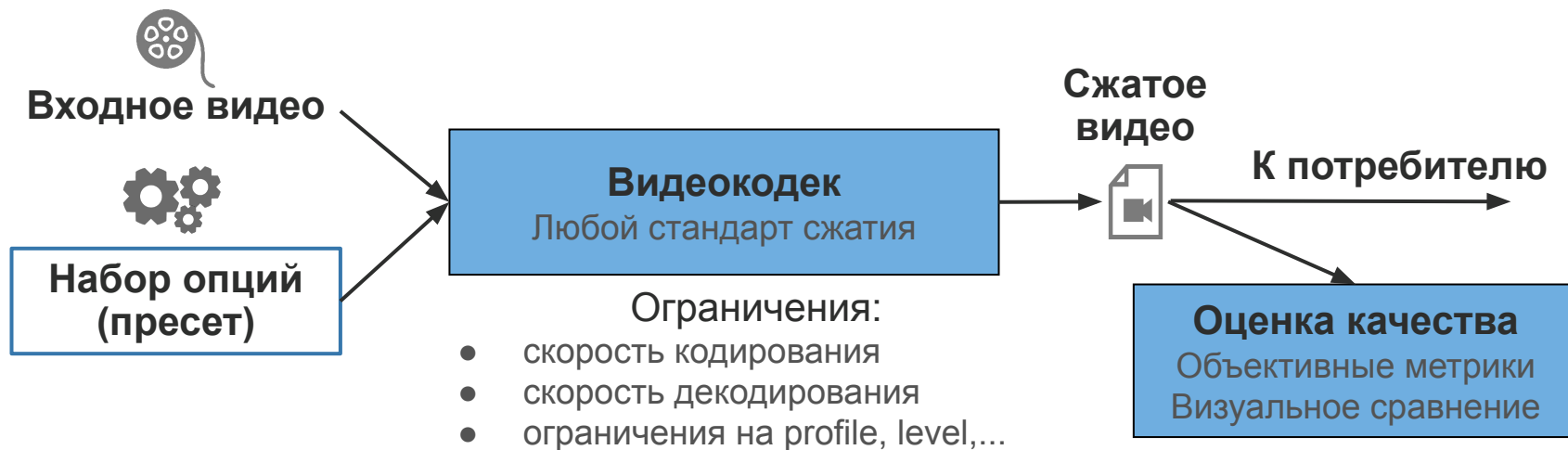
Самая известная в мире серия сравнений видеокодеков
(программных, аппаратно-ускоренных FPGA/GPU и облачных сервисов)
http://compression.ru/video/codec_comparison/index.html

Разработчики сами присылают настройки кодирования

- **18 лет** проводим сравнения кодеков
- Выпущено **40+ отчетов** (45000+ графиков, 3000+ страниц)
- **240+ кодеков** проанализировано
- Более **40 участников** (Google, Intel, Huawei, AMD и др.)



Оптимизация опций запуска видеокodeка: выбор пресета

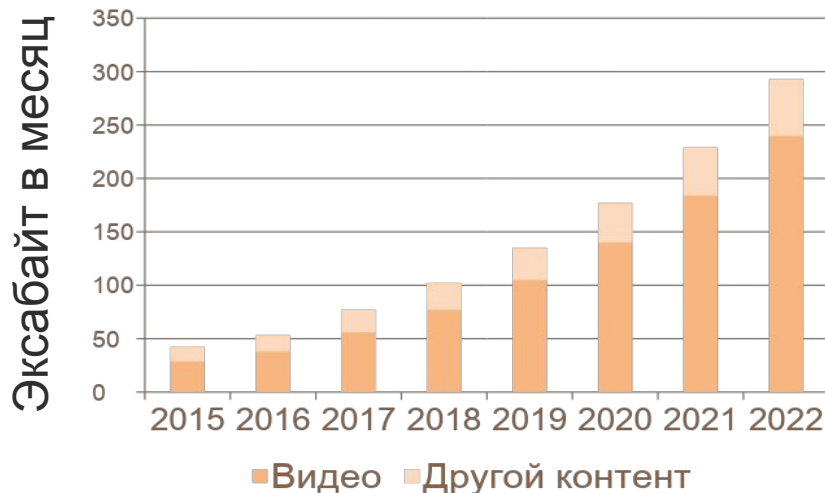


Наша основная задача: уменьшить размер видео с помощью оптимизации процесса кодирования для снижения затрат на его хранение и передачу по сети при сохранении визуального качества

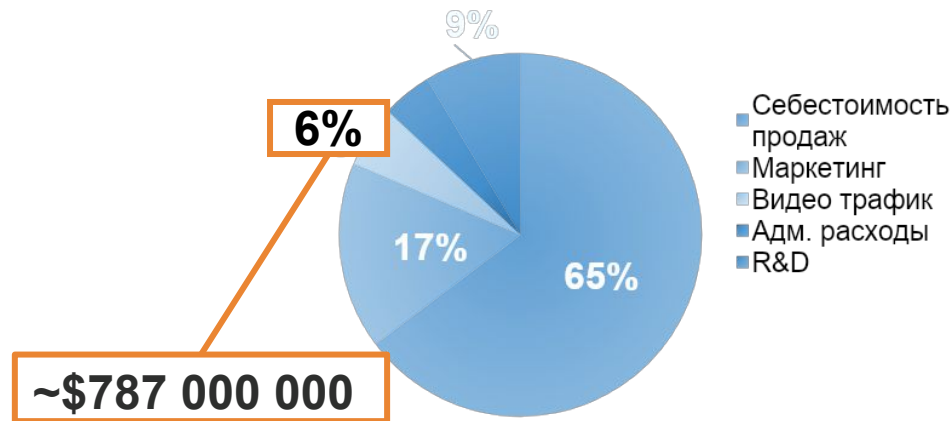
Оптимизация опций запуска видеокодека: выбор пресета

Задача супер актуальная, речь идет о миллиардах долларов

К 2022 году видео будет занимать
82% мирового интернет-трафика*



Структура расходов Netflix (2018 г.)



* InterDigital Report 2020

https://www.interdigital.com/white_papers/the-sustainable-future-of-video-entertainment

Предыстория

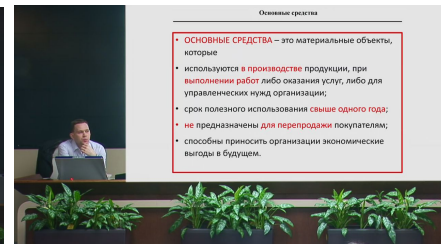
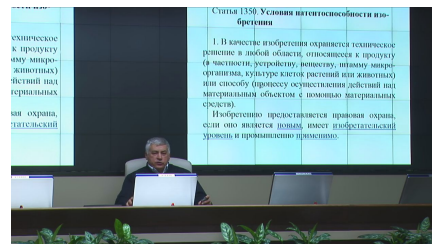
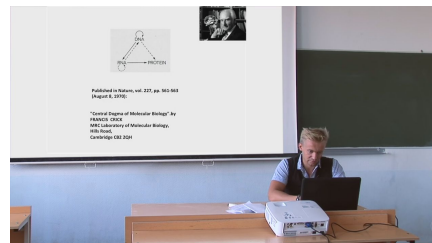
Кейс лектория

Задача: сокращение размера видео для онлайн-лектория

Наше решение: пресет для x264, который **уменьшил размер лекций на 49.28%** с тем же визуальным качеством

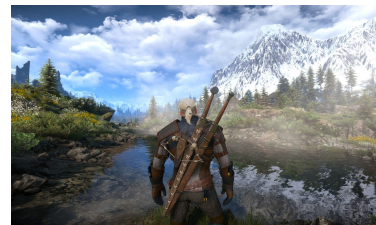
Особенности:

- **видео похоже**
- **отличаются от “типичных”**
(кодеки не были настроены на кодирование видео такого характера)



Примеры видео лектория

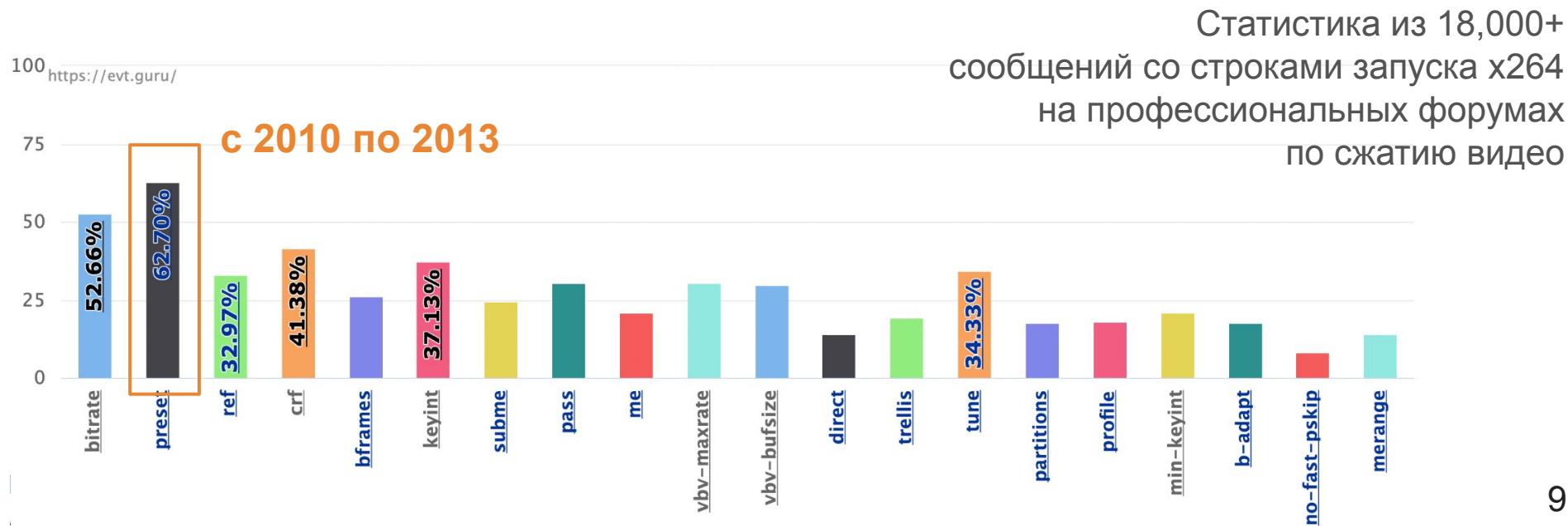
Если ваши видео похожи и специфичны (спортивный канал, мультфильмы, лекции, гейминг и т.д.), то с большой вероятностью можно сжимать их существенно лучше, чем при стандартных настройках



Но самыми широко используемыми до сих пор остаются стандартные настройки

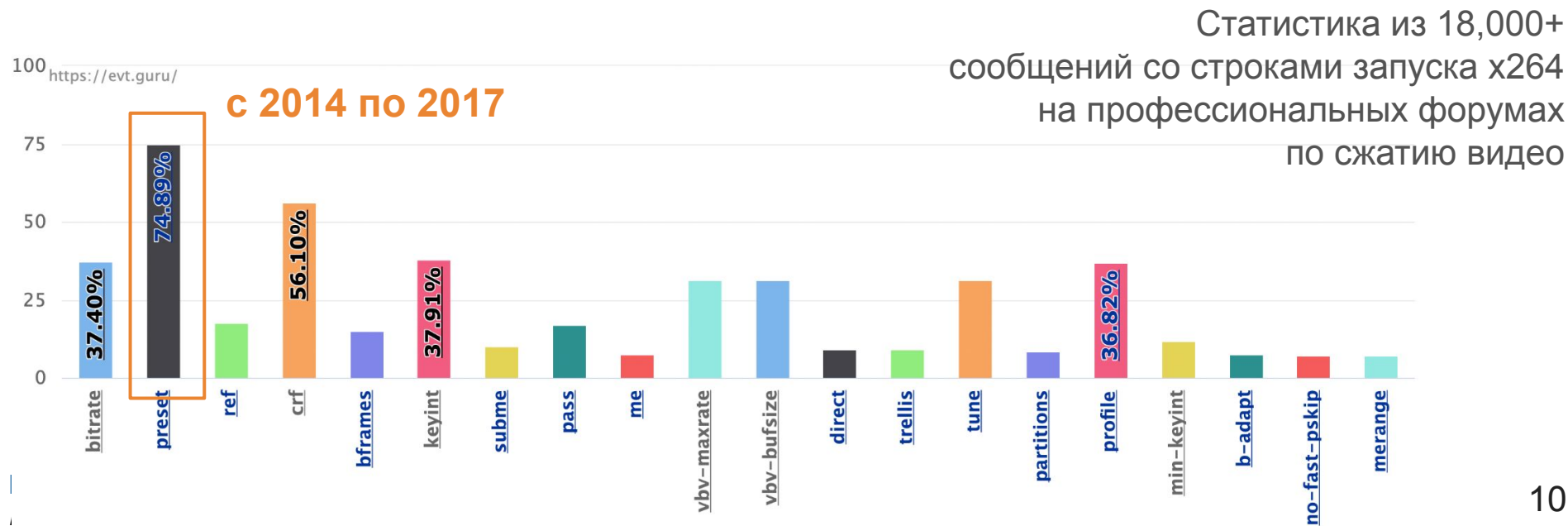
Стандартные пресеты

Стандартные пресеты (опция `--preset`) наиболее популярны у пользователей x264



Стандартные пресеты

Стандартные пресеты (опция `--preset`) наиболее популярны у пользователей x264

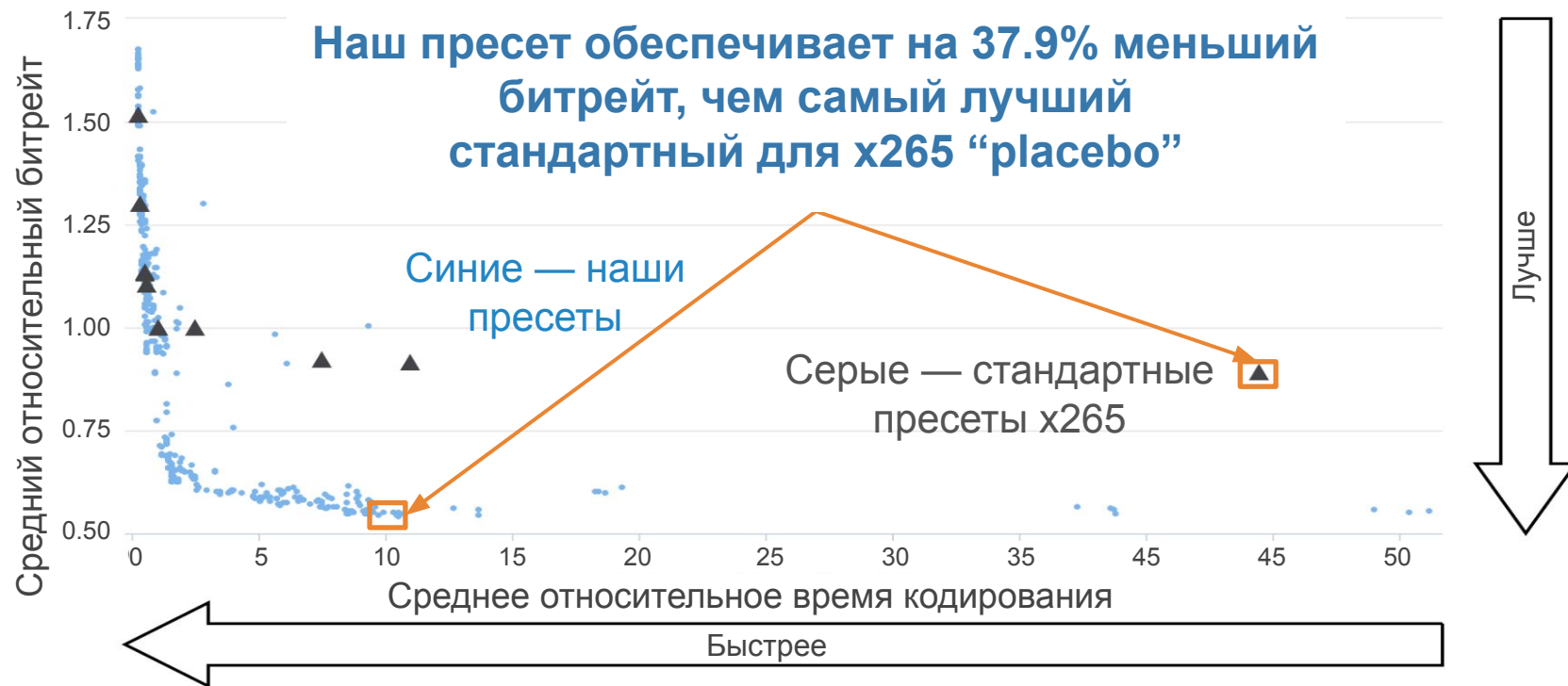


Пределы оптимизации



<https://vimeo.com/226234095>

Пределы оптимизации



Почему их пресеты были не оптимальны?

Проблемы датасетов

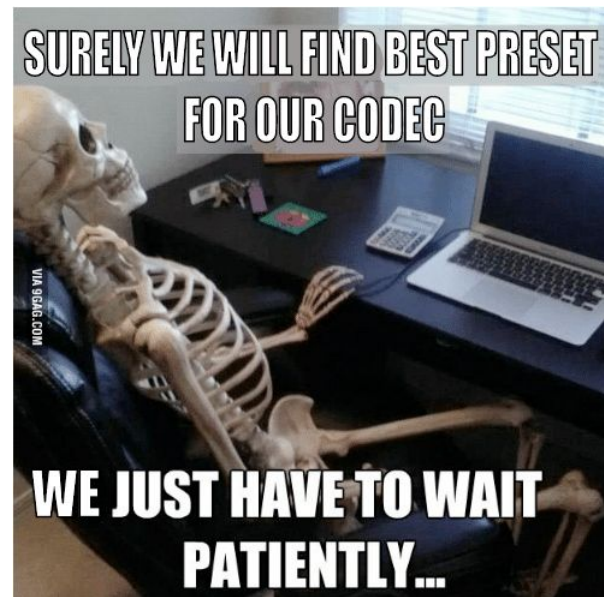
Мы проанализировали **1,000,000+** видео с сайта Vimeo



Полный перебор невозможен

Найти лучшие параметры x264 для одного **20 сек. 1080p** видео

- У x264 **45+ опций**
- Полный перебор для 1 минуты кодирования, 7 битрейтов и 3 запусков займет **$6.6 \cdot 10^{62}$ лет счета**
- С оптимизациями **$\geq 2 \cdot 10^{15}$ лет** (~500,000 возрастов Земли)



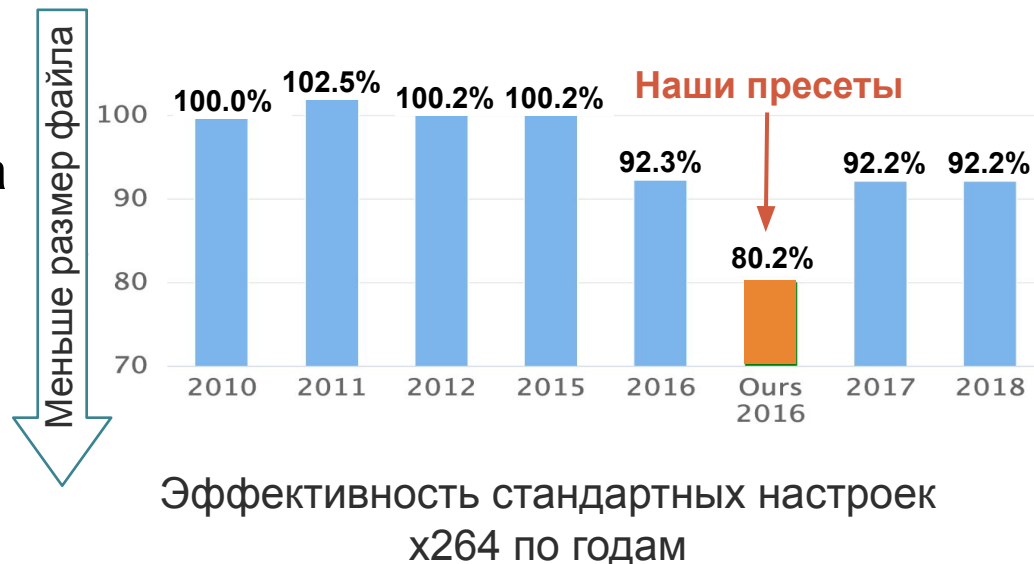
Варианты решения задачи

1. Прочитать документацию и разобраться в кодеке

Варианты решения задачи

1. Прочитать документацию и разобраться в кодеке

- Но некоторые кодеки имеют 700+ опций
- Даже разработчики кодека присылают в сравнения пресеты, далекие от оптимальных



Варианты решения задачи

1. Прочитать документацию и разобраться в кодеке
2. Обратиться к помощи зала — форумы и т.д.

Варианты решения задачи

1. Прочитать документацию и разобраться в кодеке

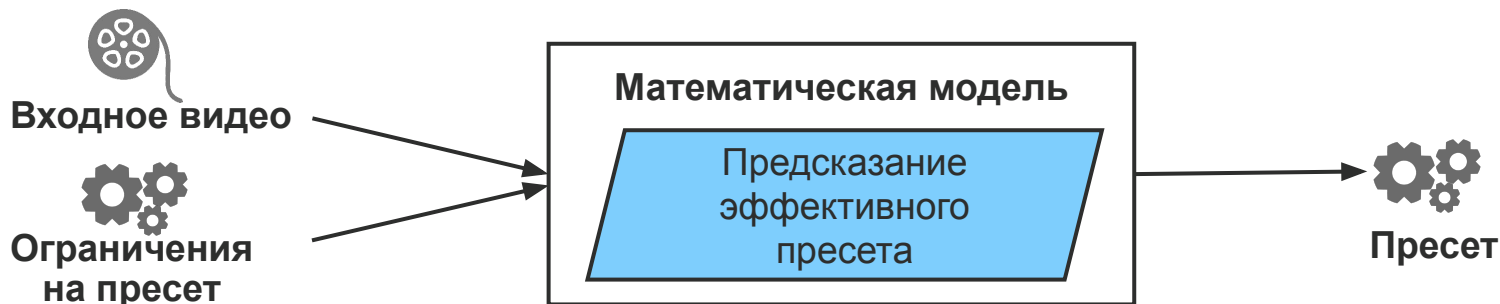
2. Обратиться к помощи зала — форумы и т.д.

Скорее всего, то, что вы найдете, все равно будет неоптимально, либо будет оптимально... не для ваших видео

Варианты решения задачи

1. Прочитать документацию и разобраться в кодеке
2. Обратиться к помощи зала — форумы и т.д.
3. Нанять профессионалов

Задача оптимизации внешних параметров кодека (тюнинг кодека)



Наш метод решения состоит из:

- модели предсказания сжимаемости видео по дескриптору его пространственно-временной сложности
- алгоритма и модели поиска Парето-оптимальных пресетов для заданного видео
- модели предсказания пресетов для заданного видео

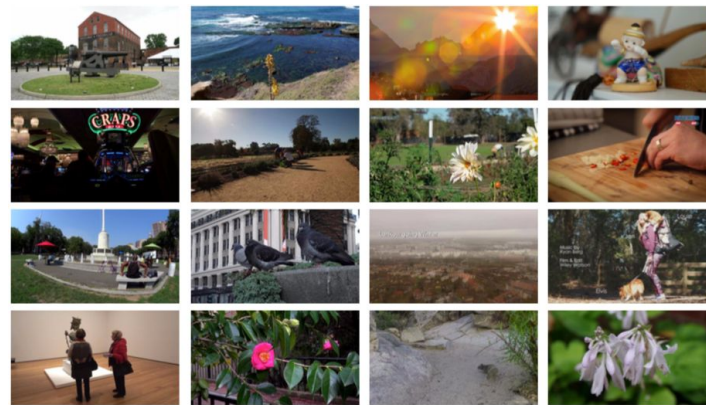
Создание баз запусков

Одним из самых трудозатратных этапов является создание базы запусков:

- набор видео
- набор пресетов
- результат кодирования этих видео этими пресетами (качество видео, скорость кодирования, размер файла)

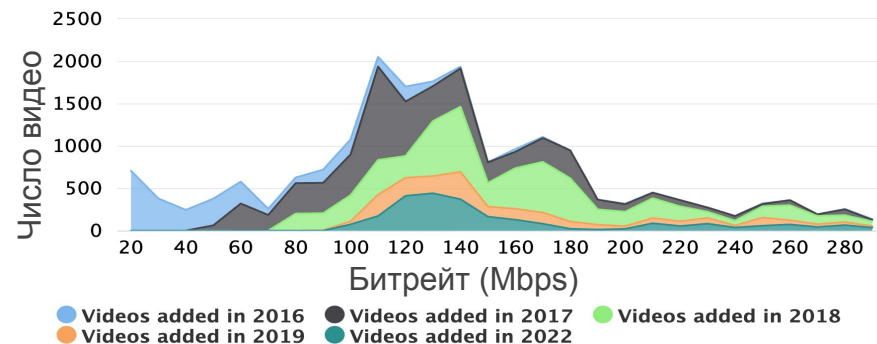
База содержит результаты для:

- > 25 тысяч видео сэмплов
- > 1 млн пресетов
- 9 кодеков



Как мы выбирали видео для построения моделей

- Видео выбраны после анализа **1,080,000+** страниц FullHD & 4K видео с Vimeo.com
- Высокий битрейт
- Видео созданы реальными пользователями, в основном профессиональный контент



Распределение битрейта в видеокolleкции

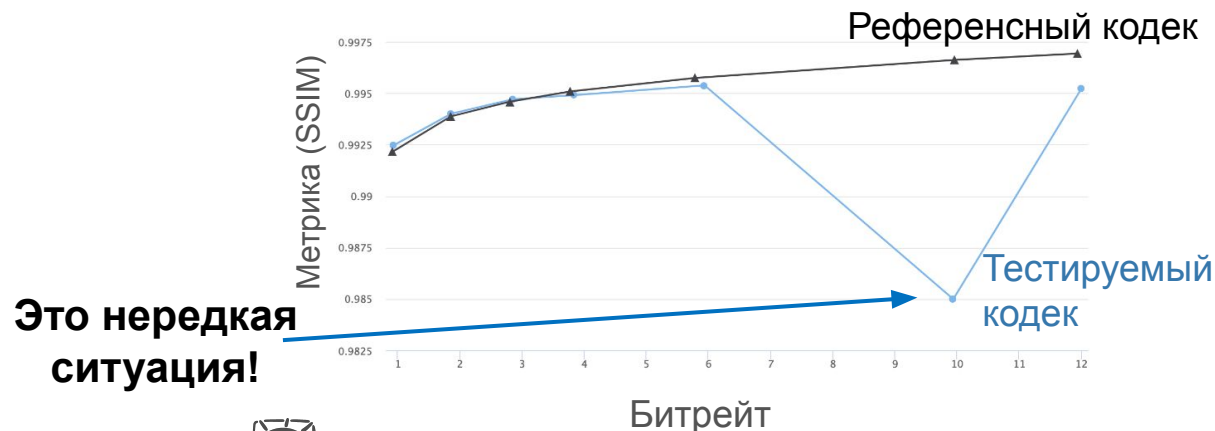
Количество отобранных видео в нашей базе, накопленный итог:

Год	# FullHD видео	# FullHD сэмплов	# 4K видео	# 4K видео	Всего видео	Всего сэмплов
2016	3	7	882	2902	885	2909
2019	4945	12402	2091	6016	7036	18418
2022	7379	19546	2091	6016	9470	25562

Некоторые сложности в построении базы и моделей

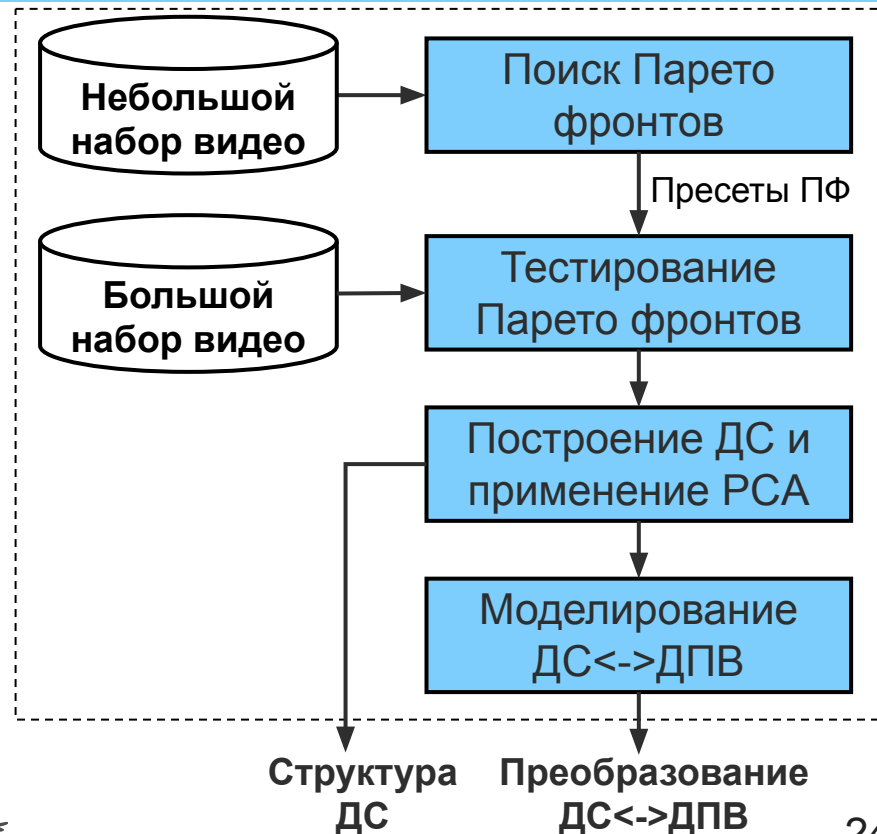
- Даже небольшие изменения в пресете влияют на результат
- Разные версии кодека могут сильно различаться
- Кодеки падают:
 - x264 – 2%, x265 – 5%
 - Кодек В – 30%
- Кодеки ведут себя непредсказуемо

```
x265 [info]: Thread pool created using 2 threads
x265 [info]: Slices : 1
x265 [info]: frame threads / pool features : 1 / wpp
x265 [info]: Coding QT: max CU size, min CU size : 32 / 8
x265 [info]: Residual QT: max TU size, max depth : 32 / 1
x265 [info]: ME / range / subpel / merge : hex / 6
x265 [info]: Keyframe min / max / scenecut : disable
x265 [info]: Lookahead / bframes / badapt : 10 / 3
x265 [info]: b-pyramid / weightp / weightb : 1 / 0 /
x265 [info]: References / ref-limit cu / depth : 2 / off
x265 [info]: AQ: mode / str / qg-size / cu-tree : 2 / 0.8
x265 [info]: Rate Control / qCompress : ABR-200
x265 [info]: tools: rd=2 early-skip rskip splitrd-skip nr-
x265 [info]: tools: tmvp fast-intra lslices=4 deblock sao
Segmentation fault
```

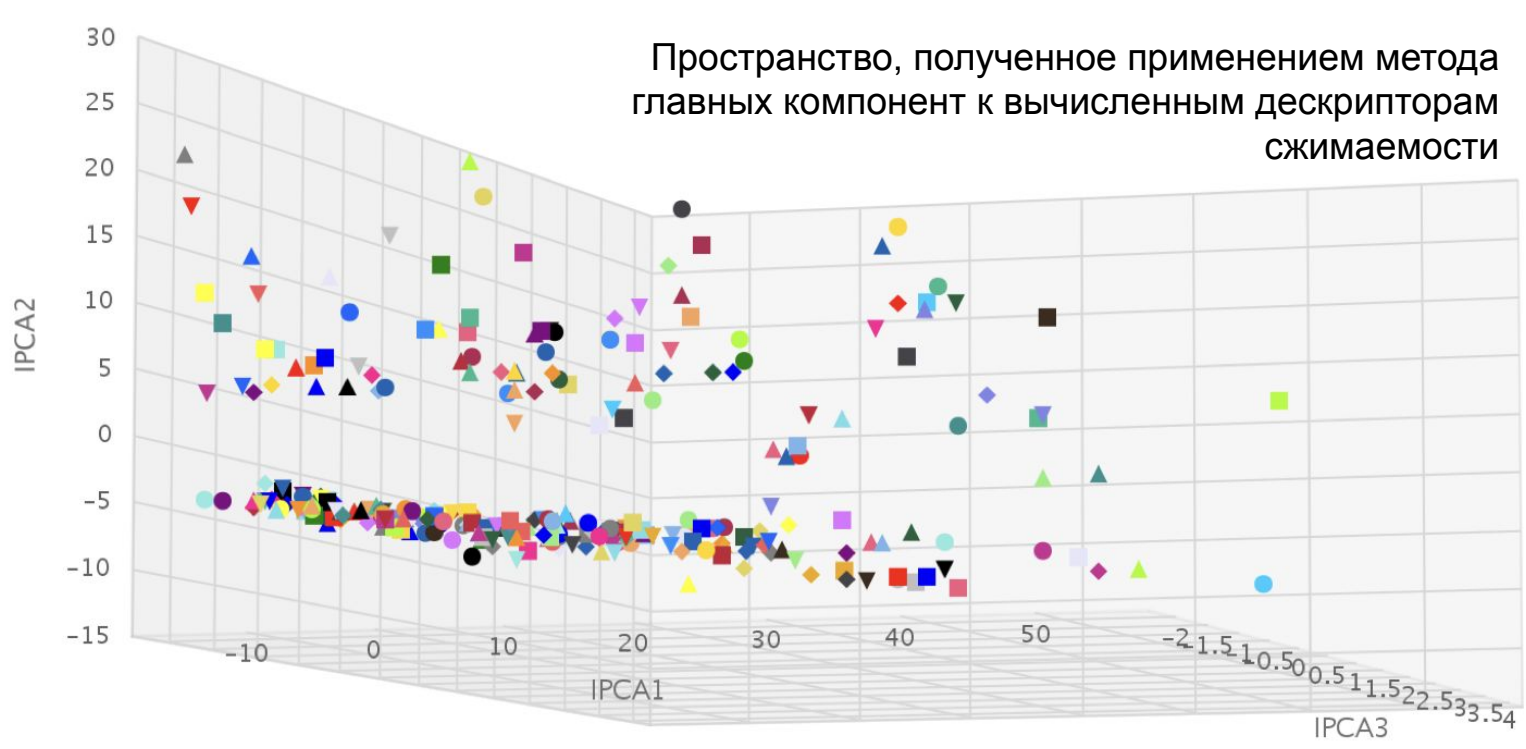


Построение дескрипторов сжимаемости видео

- Дескриптор сжимаемости (ДС) необходим для предсказания сжимаемости видео конкретным кодеком при разных параметрах запуска
- Построение ДС дорогое → нужен его быстрый аналог
- Мы обучили преобразование из быстрых пространственно-временных признаков видео (ДПВ) в урезанный ДС

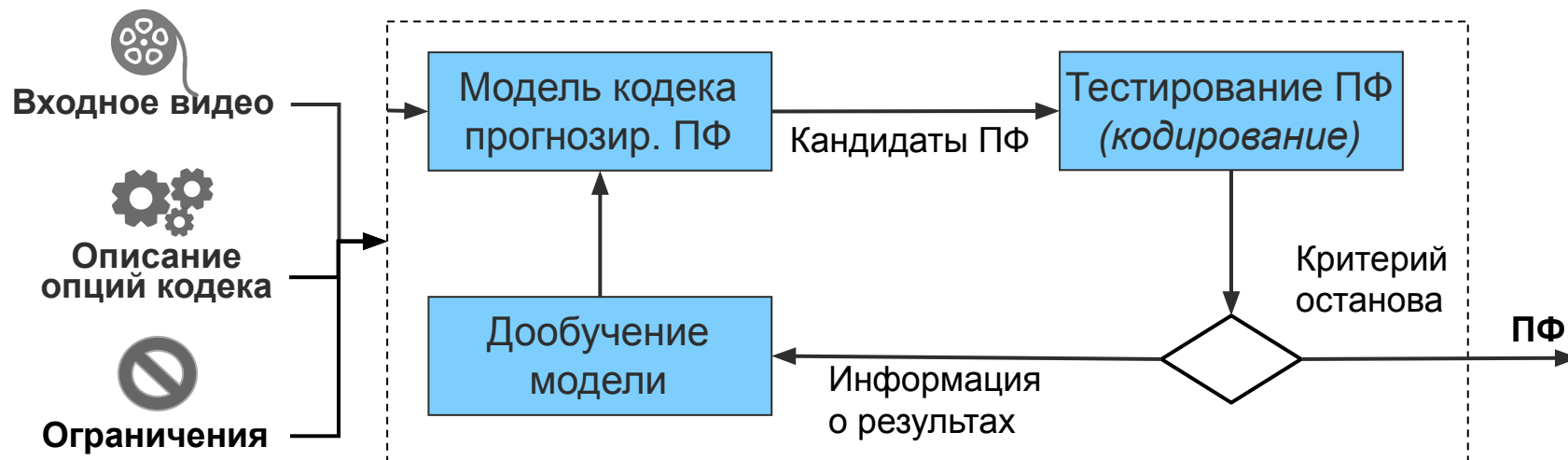


Построение дескрипторов сжимаемости видео



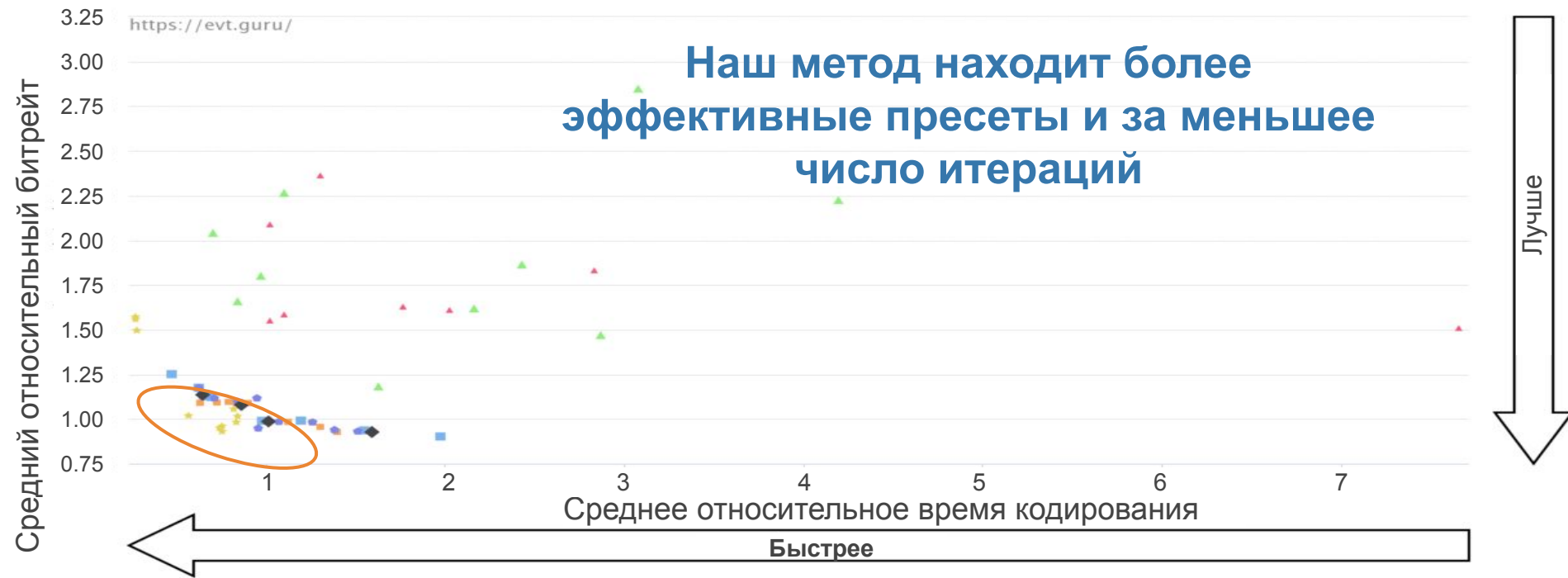
Каждому видео соответствует точка в 1372-мерном пространстве.

Поиск Парето-оптимальных пресетов



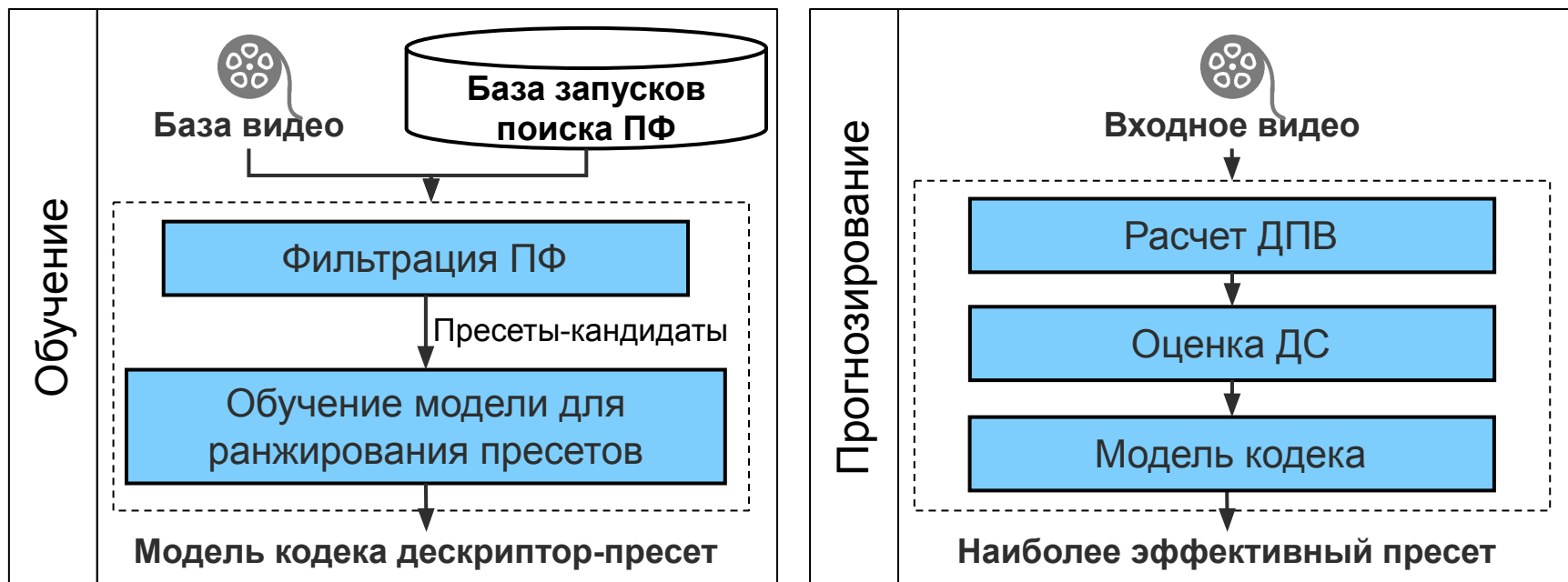
- На входе принимается:
 - видео
 - кодек и описание его опций
 - ограничения на требуемый ПФ
- На выходе:
 - аппроксимация ПФ по относительной скорости и среднему битрейту кодирования

Существовавшие алгоритмы оптимизации плохо решали задачу



Предсказание пресета для видео

Небольшая быстрая модель для выбора наилучшего пресета для кодирования видео (по аппроксимации его дескриптора сжимаемости)



Резюме по нашему решению

Изначально набор видео и пространство возможных пресетов слишком велики, чтобы собрать данные для обучения модели за разумное время

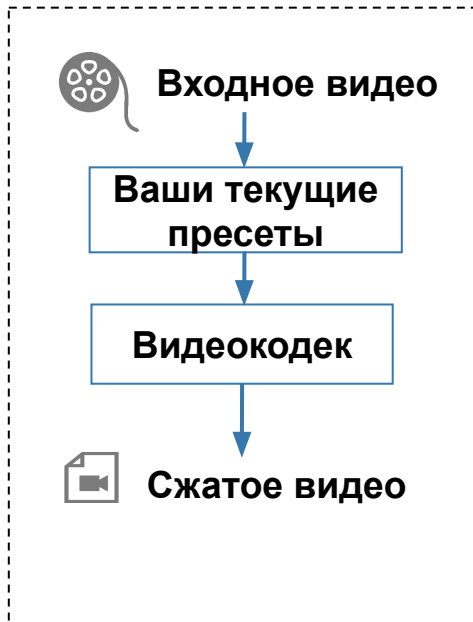
Варианты решения, которые мы использовали:

- Сузили видео датасет, нашли короткие репрезентативные дескрипторы для предсказания сжимаемости видео конкретным кодеком
- Построили модели кодеков, научились адаптировать модели под новые видео и новые версии кодека

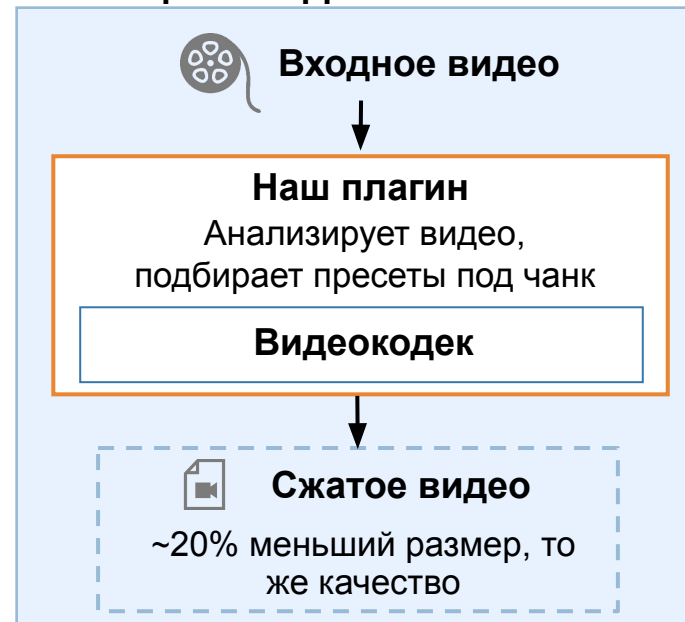
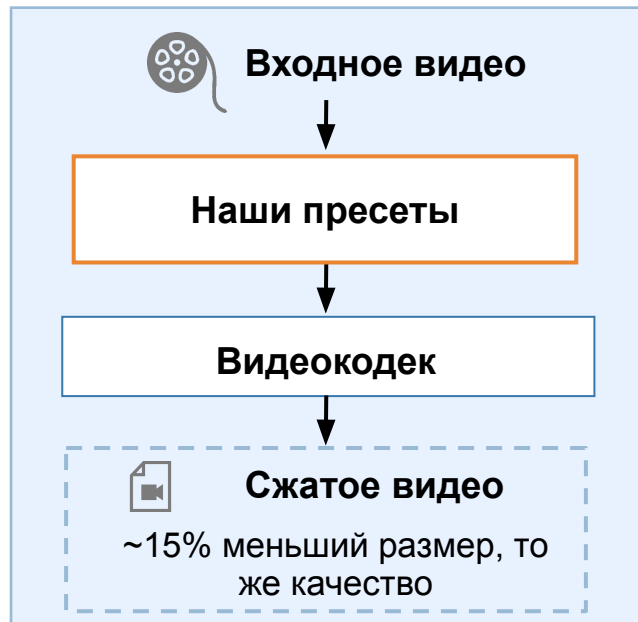
... и все равно потратили 300 машино-лет на кодирование и обучение моделей (x264, x265, VP9, OpenH264 и еще 5 коммерческих кодеков)

Как это применять

Обычный пайплайн



Пайплайны с оптимизацией кодека

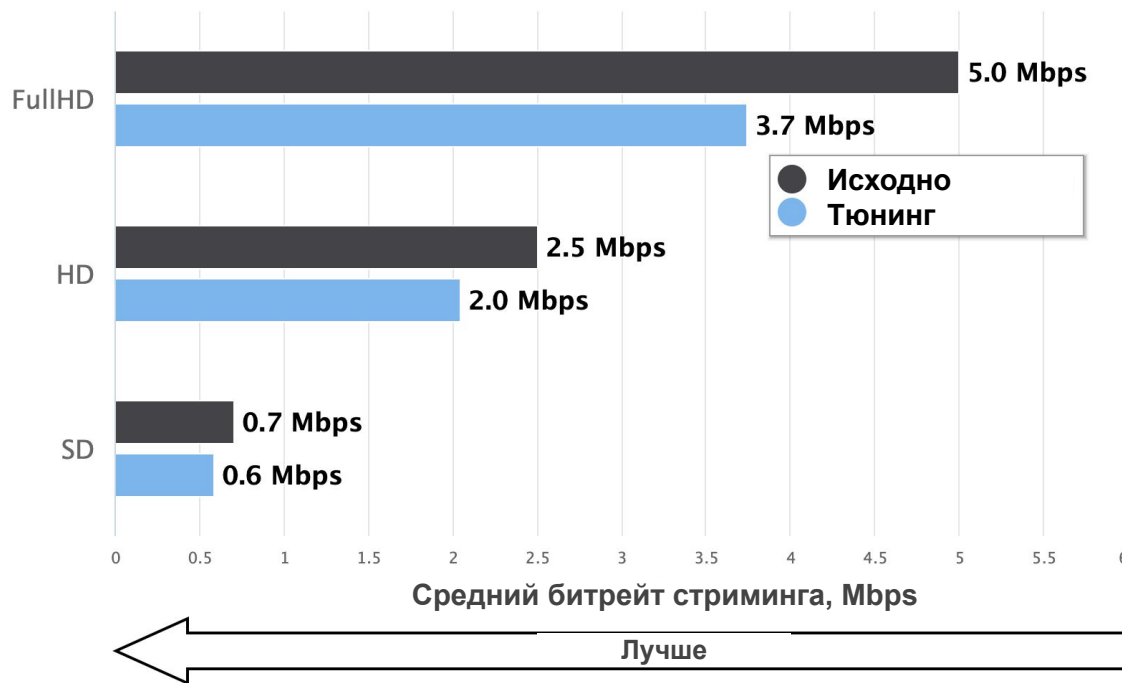


Как мы уже помогли компаниям

Кейс стримингового сервиса

Оптимизировали opnh264 и снизили битрейт в среднем на **19.5%** с тем же уровнем качества видео

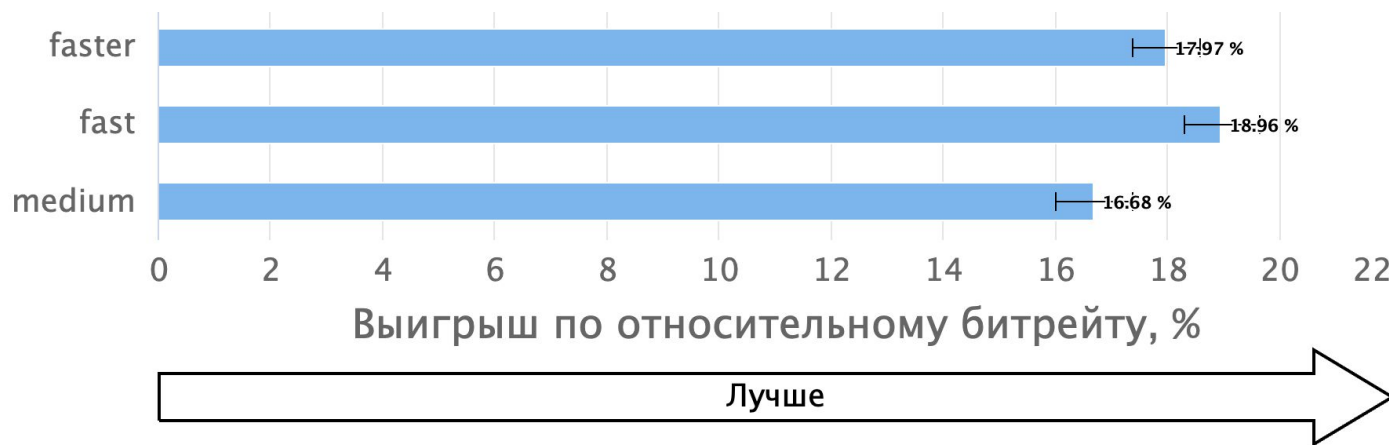
Результаты тюнинга для
трех разрешений видео



Как мы уже помогли компаниям

Кейс корпоративного видеохостинга

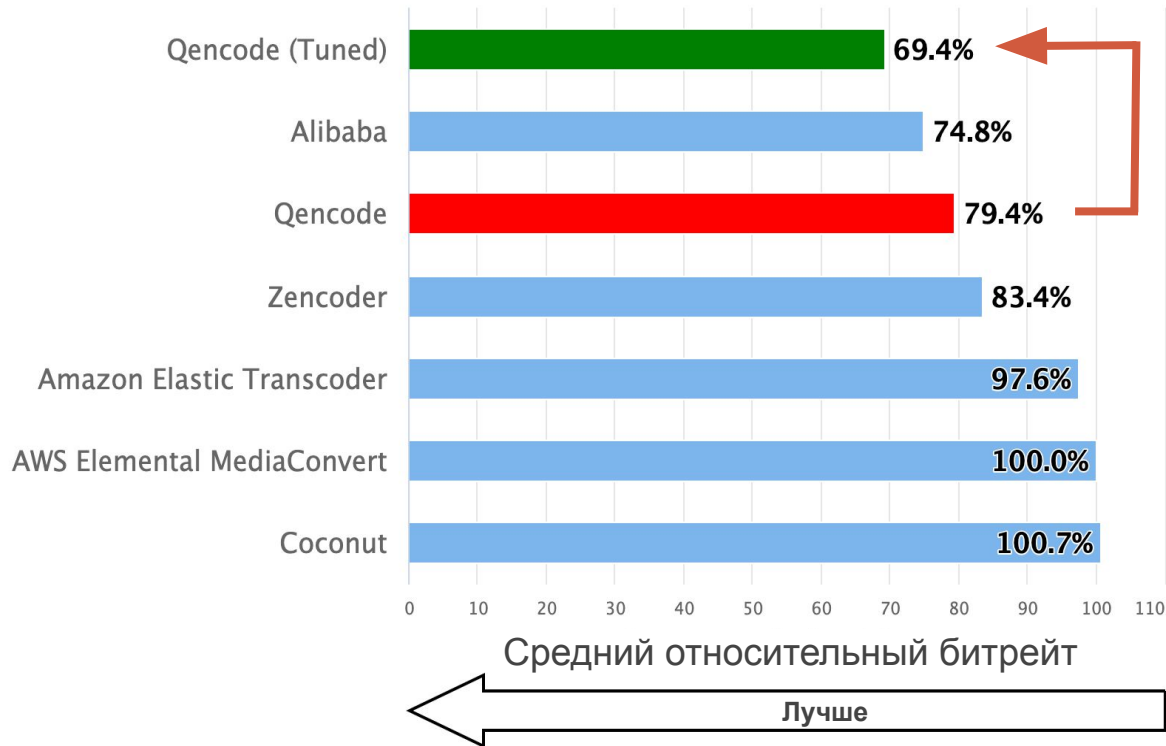
- “Корпоративный Vimeo” (лекции, презентации, реклама, спорт)
- Выигрыш 17-19% по относительному битрейту при сохранении качества и не меньшей скорости кодирования относительно пресетов x264, разрешение FullHD



Как мы уже помогли компаниям

Кейс сервиса онлайн-транскодинга

- Сложность с оптимизацией онлайн-сервисов: доступно мало опций x264 для варьирования
- Наши пресеты снизили битрейт для Qencode/x264 в среднем на 12%, что дало лучшее качество в MSU Cloud Transcoding comparison 2019

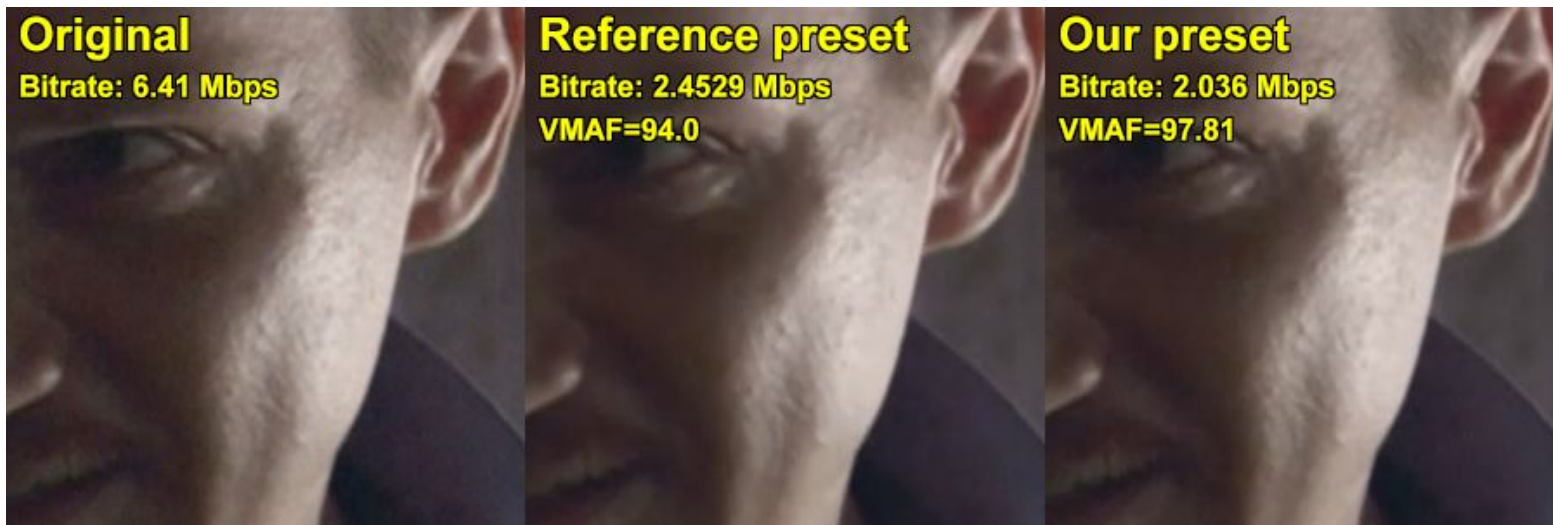


Как мы уже помогли компаниям

Кейс телеканала

Задача: улучшить качество видео, baseline profile

Мы повысили качество, подобрав пресет x264 с качеством, достижение которого потребовало бы повышения битрейта на 12% для GOP 250 и на 45% для GOP 75



Как мы уже помогли компаниям

Кейс коммерческих видеокодеков

Задача: снизить битрейт и повысить визуальное качество, сохранить значение метрики VMAF/VMAF NEG

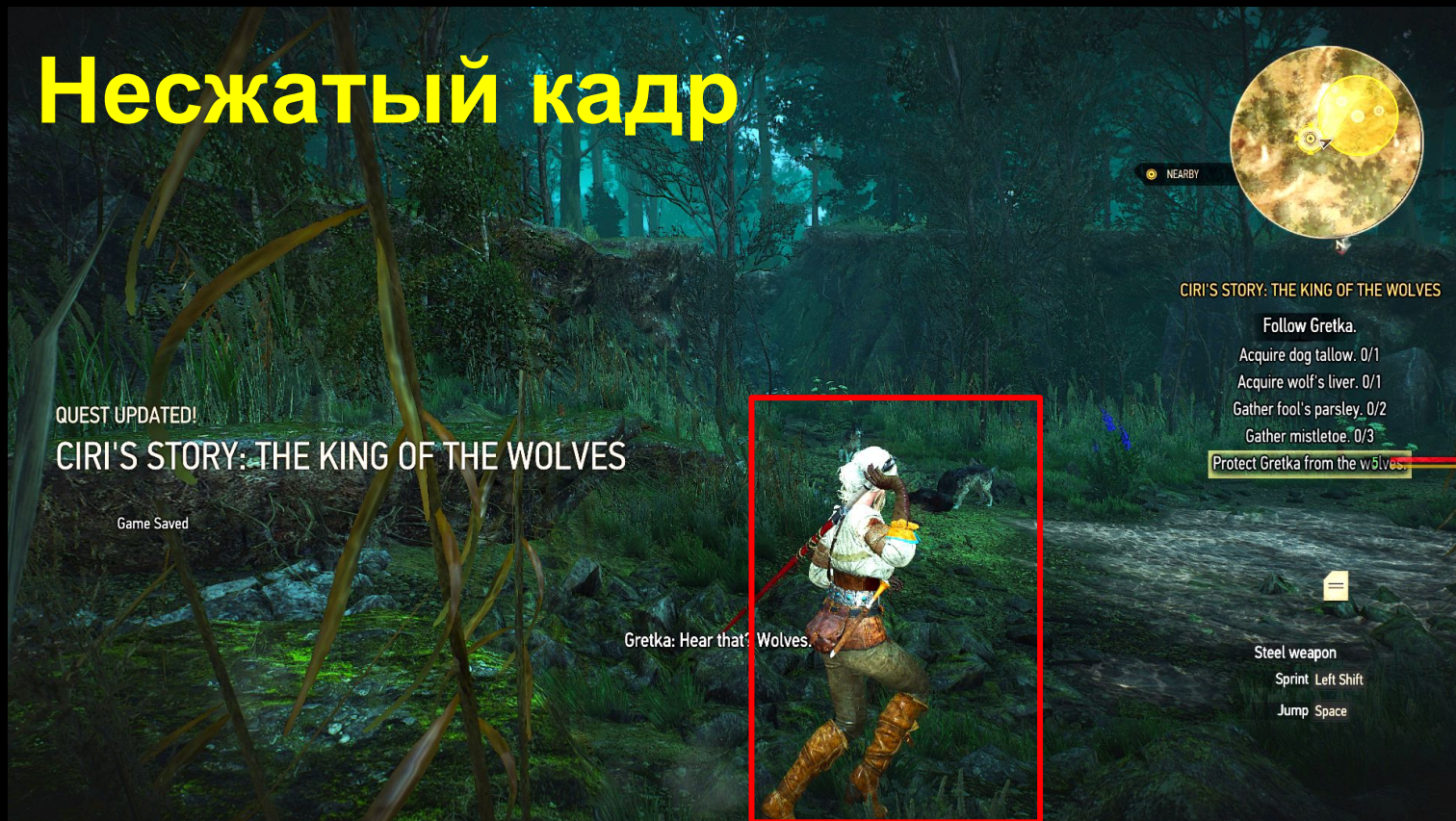
Наше решение: пресеты, которые

- для кодека стандарта H264/AVC достигли снижения битрейта на 5.29% при использовании метрики качества VMAF NEG
- для кодека стандарта H.265/HEVC достигли снижения битрейта на 11.75% при использовании VMAF NEG

Как мы уже помогли компаниям

Кейс коммерческих видеокодеков

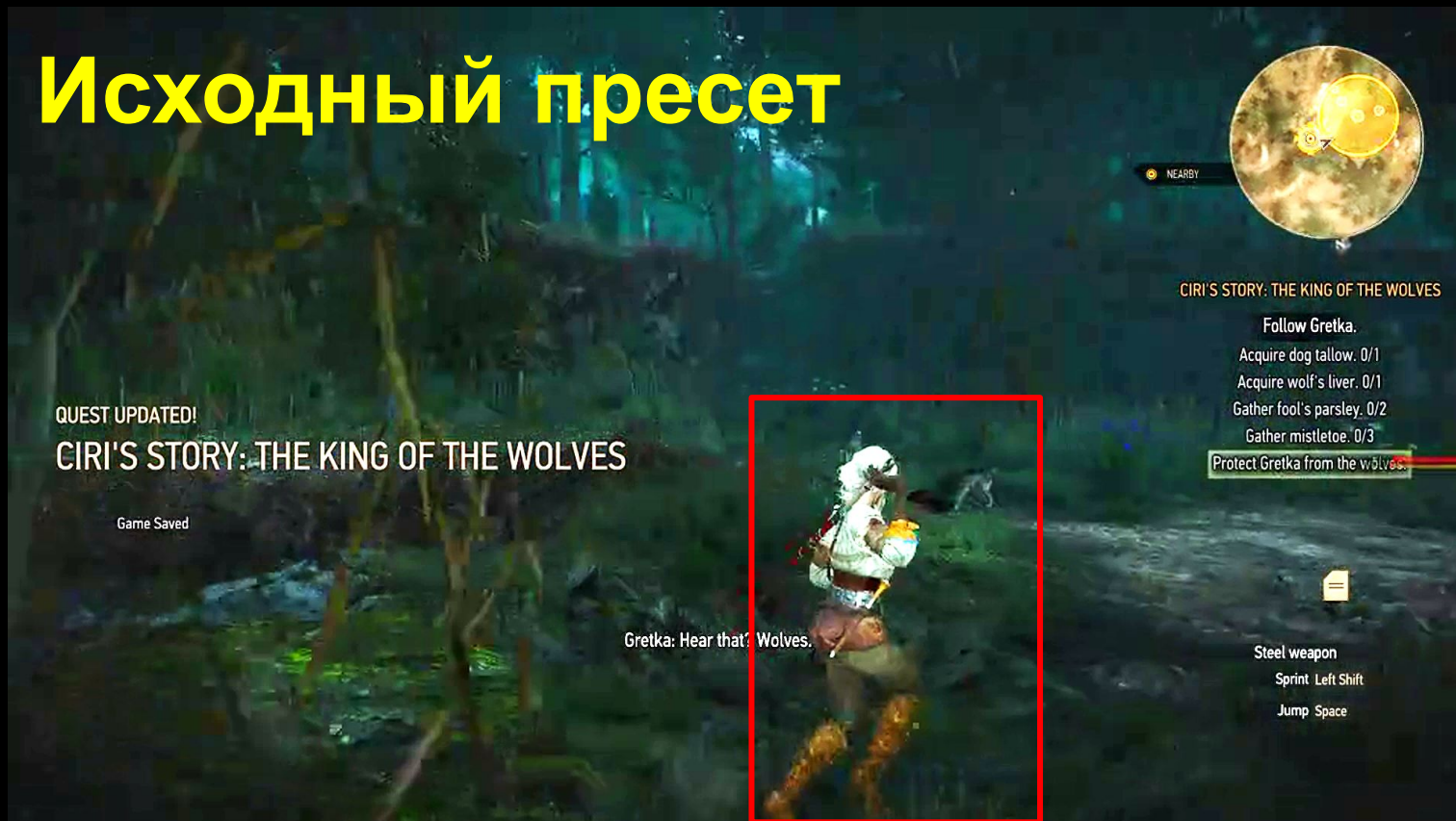
Несжатый кадр



Как мы уже помогли компаниям

Кейс коммерческих видеокодеков

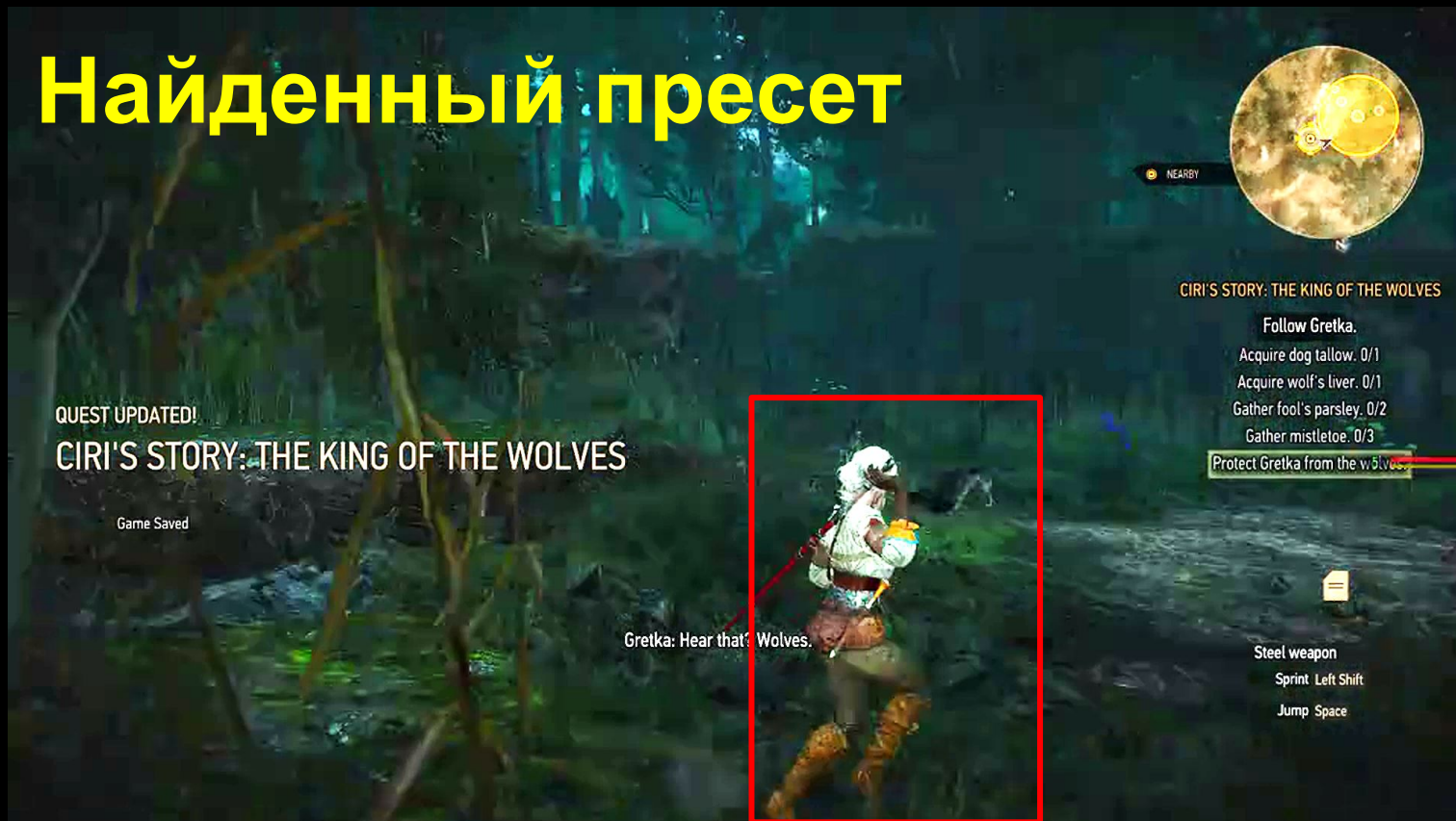
Исходный пресет



Как мы уже помогли компаниям

Кейс коммерческих видеокодеков

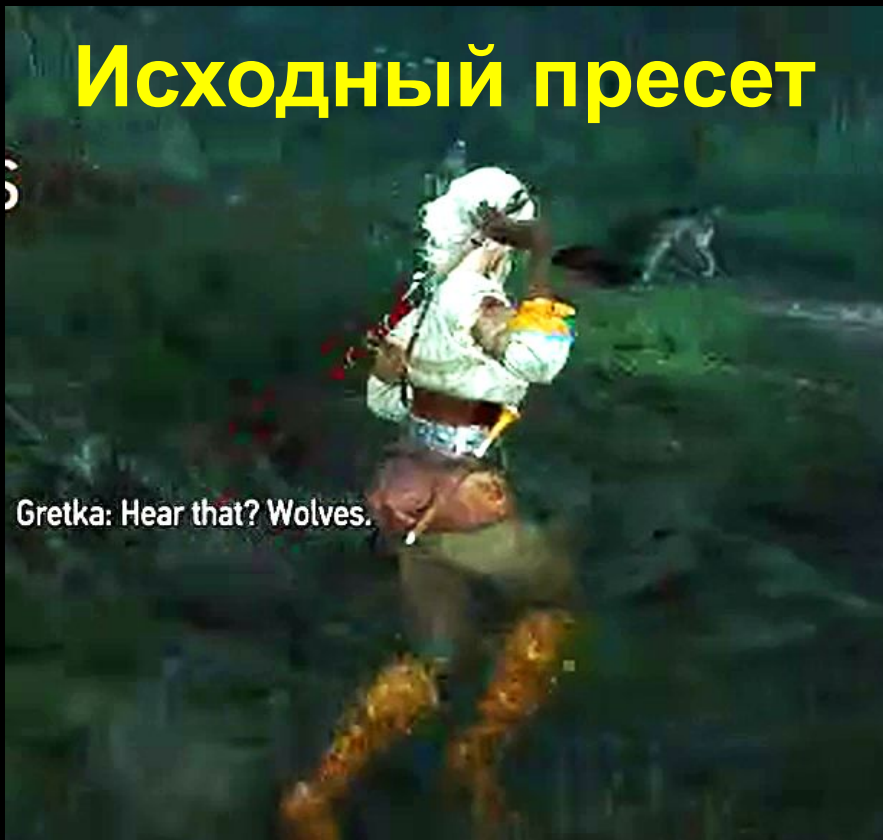
Найденный пресет



Как мы уже помогли компаниям

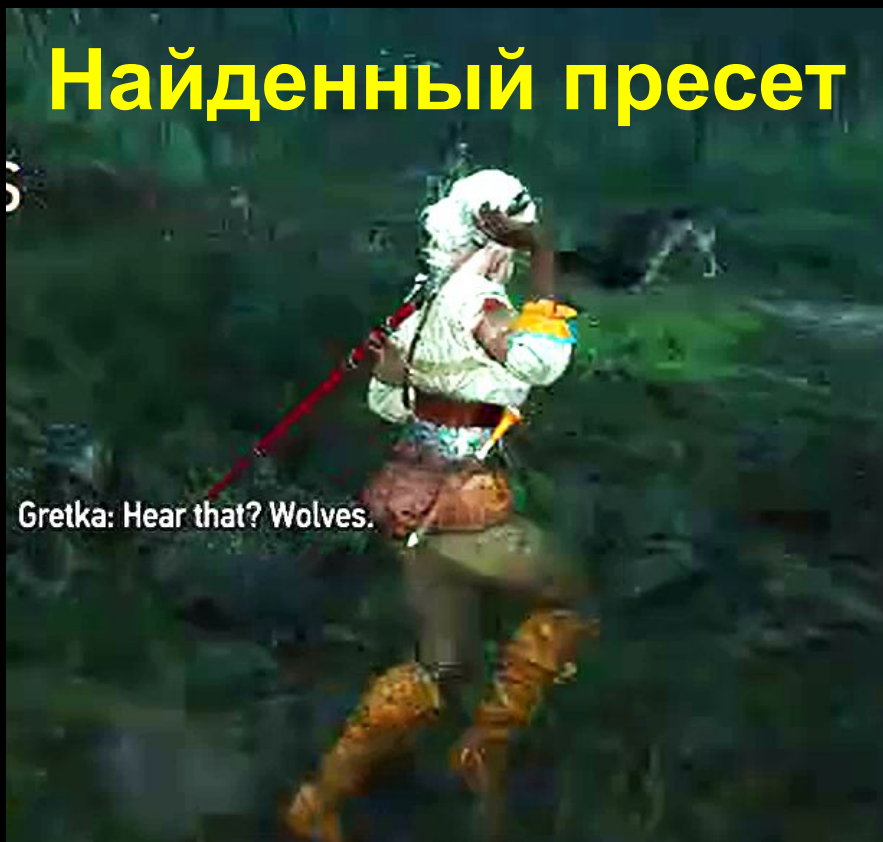
Кейс коммерческих видеокодеков

Исходный пресет



Как мы уже помогли компаниям

Кейс коммерческих видеокодеков



Варианты решения задачи

1. Читать документацию и разобраться в кодеке
2. Обратиться к помощи зала - форумы и т.д.
3. Нанять профессионалов
4. Разработать самостоятельно
 - Наше решение потребовало нескольких лет исследований и разработки. Ушло больше 2 лет для получения первых хороших результатов. Для текущего подхода потрачено порядка 300 машино-лет для построения моделей нескольких кодеков.
 - Если вы планируете аналогичную разработку, то разумно ожидать такой же порядок затрат.

Заключение

- Есть очень большой резерв оптимизации работы видеокодеков через умный подбор внешних параметров
- Максимальным пределом оптимизации для невырожденных случаев мы видим как уменьшение битрейта на 40% при том же качестве и примерно той же скорости кодирования
- Для x264 в среднем можно получить выигрыш в 15-20%
- Для x265 можно получить выигрыш в 10-15%

Заключение

- Задачу оптимизации пресета под набор видео можно решать быстро, имея эффективно работающие модель кодека и метод построения дескриптора видео
- Наше решение потребовало нескольких лет исследований и разработки и порядка 300 машино-лет для построения датасета и моделей (x264, x265 и еще 7 кодеков)
- Если вы планируете аналогичную разработку, то разумно ожидать такой же порядок затрат
- Мы предлагаем как сервис по подбору пресетов под ваши видео, так и плагин для автоматического выбора