

## **Методы для объективной оценки качества видеокодеков по сжатым ими видеопоследовательностям**

Дмитрий Ватолин, Александр Паршин

Лаборатория Компьютерной Графики и Мультимедиа, каф. АСВК, ВМиК, МГУ

{dmitriy, aparshin}@graphics.cs.msu.ru

### **Введение**

В последнее время из-за развития индустрии цифрового видео значительно возросла актуальность хороших алгоритмов сжатия видео. Сложность современных стандартов сжатия видео также возросла [1, 2], что значительно повысило качество сжатия видеопоследовательностей, но, одновременно с этим, усложнило процессы разработки и оптимизации кодеков.

Большинство сравнений видеокодеков, доступных сейчас в Интернете, являются субъективными, причём, как правило, в качестве главного эксперта выступает сам автор сравнения [3, 4]. Такая ситуация сложилась прежде всего потому, что проведение полноценное субъективное тестирование требует довольно больших ресурсов [5], а методы всестороннего объективного тестирования всё ещё не достаточно разработаны. Однако в последнее время в области объективного тестирования начали происходить некоторые улучшения: появляются новые метрики, более адекватно отражающие восприятие видеопотока человеком [6, 7], появился новый инструментарий.

В данной работе представлены некоторые новые методы объективного анализа работы кодека с использованием только закодированных последовательностей. При этом, даже не используя знания о внутренней структуре кодека, можно делать некоторые выводы относительно различных аспектов работы кодека и количестве определённых артефактов, возникающих при его работе. Авторы предлагают новые методы анализа результатов видеокодеков, а также более удачные методы визуализации полученных данных. Использование этих методов совместно с предлагаемой методологией тестирования позволило успешно завершить за последний год несколько сравнений лидирующих кодеков новых стандартов [8, 9, 10].

### **Объективные методы оценки кодеков**

Простейшим методом оценки эффективности видеокодека по закодированным им последовательностям является построение так называемой RD (Rate – Distortion) кривой для данной последовательности. Для построения этой кривой выбранная

последовательность сжимается с различными заданными значениями битрейта, и замеряется качество полученных после кодирования последовательностей (используя одну из выбранных метрик качества сжатого видео). По полученным точкам как раз и строится RD кривая. Этот метод применяется очень давно и является одним из самых распространённых при разработке кодеков. Однако он обладает рядом недостатков: не учитывает качество кодирования отдельных кадров из-за усреднения по всей последовательности, не позволяет проследить работу кодека во времени и т.д.

Мы предлагаем новый метод покадрового анализа результатов сжатия. Можно использовать покадровые значения метрики или размеры отдельных кадров, но в действительности эти величины сильно зависят друг от друга. В литературе предложено множество формул, аппроксимирующих эту зависимость [11, 12], но все они зависят от параметров видеопоследовательности. Поэтому авторы предлагают использовать в качестве аппроксимации такой зависимости экспериментально полученную RD характеристику для данного фильма, и мерить качество сжатия отдельных кадров относительно данной характеристики. Точнее, на график Битрейт/Качество наносятся точки, соответствующие отдельным кадрам, и считается расстояние со знаком до усреднённой RD кривой (если точка находится выше кривой, то знак «плюс», если ниже, то со знаком «минус»), т.е.

$$D = S \cdot \min_{i=0..n-1} \min_{t \in [0,1]} |tP_i + (1-t)P_{i+1} - P|$$

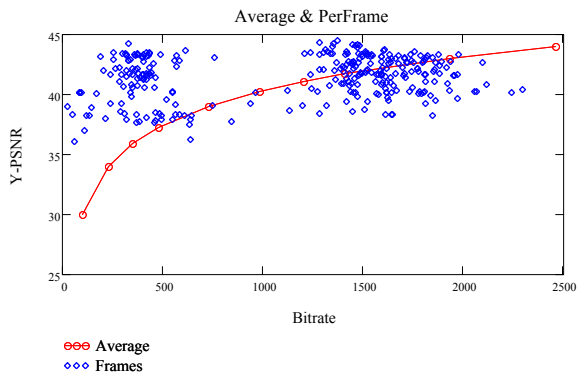
где

$P_i = \{r_i, q_i\}$  – экспериментально полученные точки усреднённой кривой

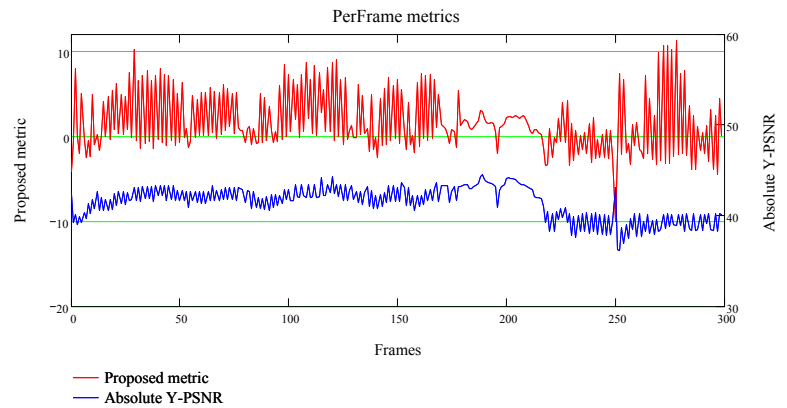
$P = \{r, q\}$  – точка, соответствующая данному кадру

$S = \pm 1$  в зависимости от того, лучше этот кадр среднего значения или хуже.

Полученные значения можно считать мерой относительного качества кадров, которые учитывают как размер, так и количество потерь при сжатии. Этот метод позволяет понять, какие из кадров были закодированы лучше, а какие хуже, учитывая как размер, так и качество.



**Рис 1. Усреднённая RD-кривая и RD-характеристики отдельных кадров**



**Рис 2. Предложенная метрика в сравнении с PSNR.**

В своём последнем сравнении мы использовали большой набор метрик. Помимо классической метрики PSNR использовались метрики VQM[6] и SSIM[7]. Кроме того, мы использовали новые метрики для выявления количества артефактов определённого вида в закодированной последовательности. При сжатии видео в рамках современных стандартов все искажения кодека вносятся на этапе квантования коэффициентов DCT (дискретное косинусное преобразование) отдельных блоков изображения (строго говоря, потери так же возникают из-за необратимости преобразования DCT, но они, как правило, не существенны). После обратного преобразования квантованных коэффициентов возникают сильные различия на границах блоков, которые называются эффектом «блочности» (blocking effect). Этот эффект особенно заметен при увеличении шага квантования, то есть при увеличении степени сжатия. С этим эффектом пытаются бороться, определяя границы блоков (возможно, с использованием информации о векторах движения) и размывая их в нужном направлении. Однако этот метод порождает слишком размытое изображение, также не очень приятное для просматривающего сжатую последовательность (так называемый эффект «размытия», blurring effect).

Мы измеряли наличие этих эффектов при помощи специальных метрик. Эти метрики были разработаны в нашей лаборатории, и используют только кадры сжатой последовательности [13]. Поэтому для сравнения необходимо использовать ещё и информацию о величине этих артефактов в исходной последовательности. Так как эффекты блочности и размытия являются одними из самых заметных во временной области, очень важно понимать, насколько сильно декодер подавляет данные артефакты. Кроме того, замеры этих артефактов можно использовать и при оптимизации работы алгоритмов управления битрейтом кодека.



Рис 3. Исходное изображение



Рис 4. Изменённое изображение

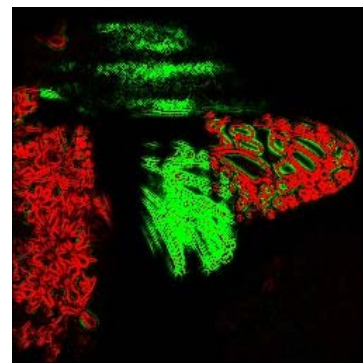


Рис 5. Визуализация метрики размытия



Рис 6. Исходное изображение

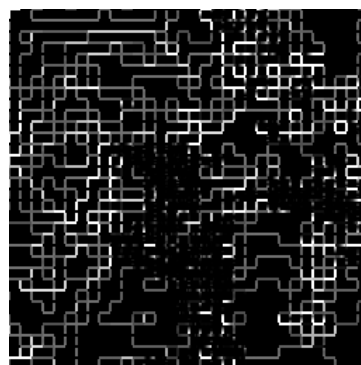


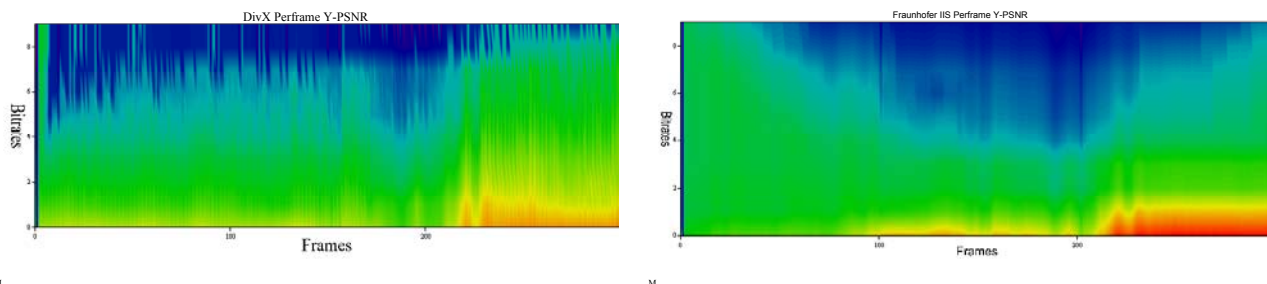
Рис 7. Визуализация метрики блочности

## Методы визуализации результатов тестирования

Главным преимуществом объективного тестирования перед субъективным является, прежде всего, возможность проанализировать большой набор последовательностей и автоматически получить множество данных. Однако отдельной проблемой является задача проанализировать полученные данные. Мы использовали двух- и трёхмерные графики для визуализации данных для повышения наглядности данных и эффективности их анализа.

Основными в проведённых тестированиях кодеков [8, 9, 10] по-прежнему являлись классические RD кривые кодеков (графики Битрейт/Качество). Однако мы применяли трёхмерные графики для анализа кадровых данных о качестве. По оси абсцисс были отложены отдельные кадры последовательности, по оси ординат – битрейт, с которым была сжата последовательность. Цветом показывается качество отдельного кадра для данного битрейта (рис 8). Используя такие довольно простые графики можно много сказать о поведении кодека, особенно об алгоритмах управления битрейтом. По ним можно проследить реакцию кодека на смену сцен, качество изначальных параметров (как

правило, просадка в начале фильма по качеству) и т.д. Более того, можно делать выводы об использовании некоторых особенностей стандартов, таких как наличие в последовательности В кадров, места вставки ключевых кадров и т.д.



**Рис 8. Примеры визуализации пок кадровых метрик. Метрика Y-PSNR, красный цвет – плохое качество, синий – хорошее.**

Если к замерам качества и битрейта добавить ещё и замеры скорости, то можно анализировать поведение кодеков в терминах соотношения качество/скорость. В своём сравнении мы рассматривали поведение кодеков для выбранной последовательности и битрейта, то есть, в осях качество/скорость каждый кодек представлялся одной точкой. Это позволяло для некоторых последовательностей определять, что один кодек лучше другого по производительности (лучше по качеству и быстрее по скорости), но для некоторых таких заключений сделать было нельзя. Кроме этих графиков были построены графики абсолютного времени кодека, графики удержания заданного битрейта, усреднённых времён и относительные графики Битрейт/Качество (см. Рис 9).

## Методология тестирования

Как уже было замечено, при проведении тестирования необходимо оперировать большими объемами, как тестовых данных, так и обрабатывать значительное количество полученных материалов тестирования. На всё это накладывается необходимость оперативно взаимодействовать с разработчиками кодеков при обнаружении в их продуктах каких-либо ошибок или недоработок. Мы так же столкнулись с этими трудностями при проведении наших первых сравнений. Для их решения была разработана определённая методология, призванная ускорить процесс проведения тестирования и минимизировать использование ручной работы на этапе замеров и построения графиков.

Для автоматизации проведения замеров была разработана система командных файлов, которая позволяла через достаточно простой интерфейс реализовать запуск кодека для кодирования выбранных последовательностей, декодирования файлов и сравнения качества декодированных последовательностей.

При этом в данной системе были учтены возможности использования разными кодеками различных вариантов форматов одной и той же последовательности. Например, в нашем последнем тестировании нам пришлось создавать три различных вида файлов – файлы .yuv в формате YV12; файлы .avi в формате YV12 и с изменённым заголовком файла (один из кодеков требовал определённого кода FourCC), а также файлы .yuv с продублированным последним кадром (один из декодеров не обрабатывал последний кадр из-за какой-то внутренней ошибки).

Была реализована возможность задавать различные настройки кодекам. Для этого было введено понятие «пресета» кодеков – их настроек, отвечающих одним и тем же условиям работы (одинаковым профилям стандартов, соизмеримой скоростью работы и т.д.). Это позволяло сравнивать кодеки в различных режимах их работы (в последнем сравнении были пресеты «Наилучшее качество», «Наибольшая скорость» и «Дополнительные замеры»). Также иногда бывает полезно изменять набор битрейтов, для которых проводится тестирование. Это позволяет локализовать некоторые проблемные места работы кодека вокруг выбранного битрейта, с большим разрешением рассмотреть одну из целевых областей применения кодеков (HDTV телевидение, видеоконференции и т.д.). Все настройки, используемые при работе кодека, собраны в профайле тестирования, который позволяет быстро переключаться между различными условиями тестирования или даже между разными тестированиями. Структура используемой нами среды тестирования показана на Рис. 10



Рис 9. Этапы процесса объективного тестирования

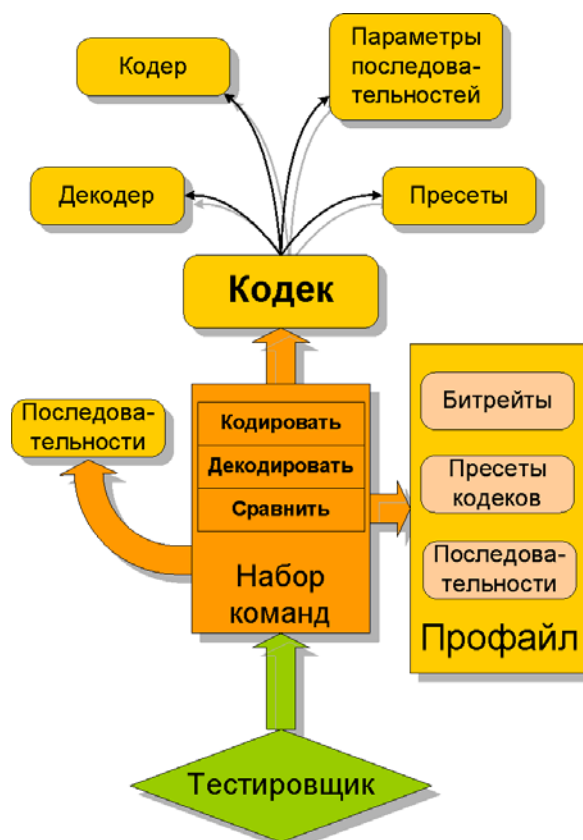


Рис 10. Структура системы автоматического тестирования

Опыт проведённых тестирований показал, что каждый кодек имеет смысл тестировать в 3 этапа (Рис 11):

1. Проверка работоспособности кодека отдельно от среды тестирования. Основная задача этого этапа – первичная проверка работы кодека и быстрая реакции на очевидные ошибки в кодеках.
2. Погружение кодека в среду тестирования. На этом этапе создаются все необходимые файлы для использования автоматической среды тестирования, и проверяется работа кодека при помощи этой среды на небольшом тестовом наборе с приблизительными замерами скорости. Это позволяет уменьшить вероятность ошибок при полном тестировании кодека.
3. Полное тестирование. Проводится автоматическое тестирование кодека на полном тестовом наборе с точными замерами скорости. Именно на этом этапе получают основные результаты тестирования. Основная задача первых двух этапов – уменьшить вероятность возникновения ошибок при полном тестировании.

Использование этого процесса позволило нам повысить скорость выявления ошибок в кодеках (в последнем тестировании мы обнаружили недоработки в большей части тестируемых кодеков), что заметно ускорило проведение самого тестирования.

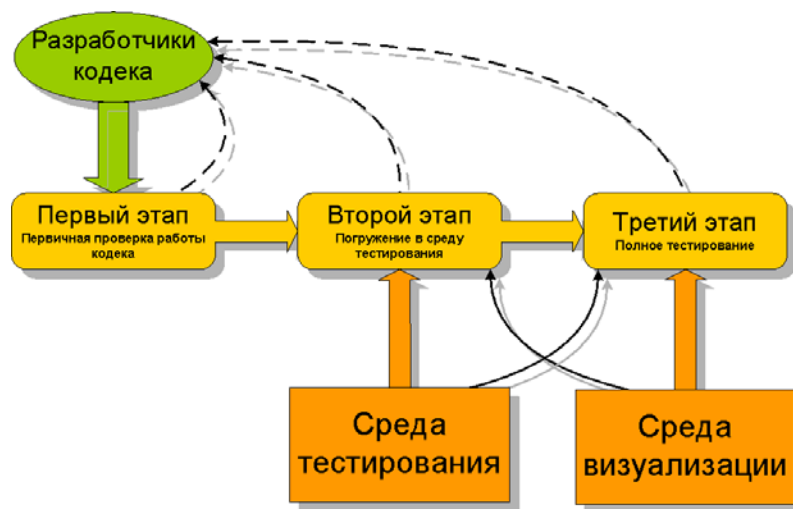


Рис 11. Этапы тестирования кодека

## Результаты

В проведённых нами объективных тестированиях кодеков мы использовали новые методы анализа кодеков по результатам их работы. Применение новых метрик оценки величины возникших при сжатии артефактов вместе с новыми подходами к оценке качества работы позволило собрать намного больше информации о поведении кодеков на тестовом наборе. Использование новых методов визуализации результатов в рамках предложенной нами методологии процесса тестирования позволило уже несколько раз удачно провести сравнения лидирующих кодеков нескольких стандартов сжатия.

## Список литературы

1. Thomas Wiegand, Gary J. Sullivan, Gisle Bjontegaard, and Ajay Luthra. Overview of the H.264 / AVC Video Coding Standard, IEEE CSVT, 2003г.
2. Mohammed Ghanbari. Standard Codecs: Image Compression to Advanced Video Coding, IEE PUBLISHING, 2003г.
3. Сравнение кодеков на сайте Doom9.org. <http://www.doom9.org/index.html?/codecs-104-1.htm>, 2004г.
4. Jason Cross. Video Codec Shootout, сайт ExtremeTech. <http://www.extremetech.com/article2/0,1697,1544886,00.asp>
5. International Telecommunication Union. Methodology for the subjective assessment of the quality of television pictures (ITU-R BT.500-11). 2002г.



6. Feng Xiao. DCT-based Video Quality Evaluation. Final Project for EE392J. 2000г.
7. Z. Wang, A. C. Bovik, H. R. Sheikh and E. P. Simoncelli. Image Quality Assessment: From Error Visibility to Structural Similarity. IEEE Transactions on Image Processing, vol. 13, no. 4. 2004г.
8. Д. Ватолин, Д Куликов, А Паршин. Сравнение MPEG-4 SP/ASP видеокодексов Сайт [www.compression.ru](http://www.compression.ru). [http://compression.ru/video/codec\\_comparison/mpeg-4.html](http://compression.ru/video/codec_comparison/mpeg-4.html), 2005г.
9. Д. Ватолин, Д Куликов, А Паршин. 1-е сравнение MPEG-4 AVC/H.264 видеокодексов. Сайт [www.compression.ru](http://www.compression.ru). [http://compression.ru/video/codec\\_comparison/mpeg-4\\_avc\\_h264.html](http://compression.ru/video/codec_comparison/mpeg-4_avc_h264.html), январь 2005.
10. Д. Ватолин, Д Куликов, А Паршин. 2-е сравнение MPEG-4 AVC/H.264 видеокодексов. Сайт [www.compression.ru](http://www.compression.ru). [http://compression.ru/video/codec\\_comparison/mpeg-4\\_avc\\_h264\\_2005.html](http://compression.ru/video/codec_comparison/mpeg-4_avc_h264_2005.html), декабрь 2005.
11. J. Ronda, M. Eckert, F. Jaureguizar, N. Garcia. Rate Control and Bit Allocation for MPEG-4., IEEE CSVT, 1999г.
12. S. Mitra, Z. He. Optimum Bit Allocation and Accurate Rate Control for Video Coding via p-Domain Source Modeling, IEEE CSVT, 2002г.
13. Д. Ватолин, А. Москвин. MSU Video Quality Measurement Tool. Сайт [www.compression.ru](http://www.compression.ru). [http://www.compression.ru/video/quality\\_measure/video\\_measurement\\_tool.html](http://www.compression.ru/video/quality_measure/video_measurement_tool.html)