

Новый метод анализа эффективности RD моделей видеокодеков и алгоритмов адаптивного подбора их параметров

Алексей Носков, Александр Паршин

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

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

Аннотация

Многие алгоритмы управления видеокодеками основываются на моделях зависимости количества вносимых искажений от коэффициента сжатия (RD моделях). В статье рассматривается вопрос методов оценки качества этих моделей, не зависящих от использующего их кодека и стандарта сжатия. Помимо непосредственно самих моделей, рассмотрены алгоритмы обновления адаптивных параметров RD моделей. Произведён анализ как ряда известных, так и нескольких предложенных авторами RD моделей, сделаны заключения относительно их эффективности и возможностей её улучшения.

Ключевые слова: Rate-Distortion модели, анализ кодеков, управление битрейтом.

1. ВВЕДЕНИЕ

1.1 RD модели в современных видеокодеках

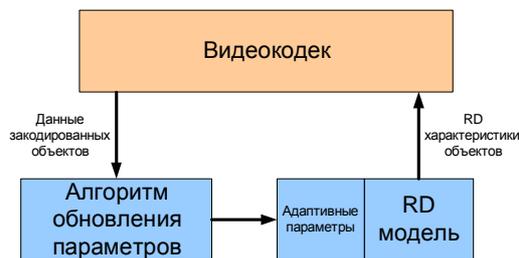


Рисунок 1: Модель использования кодеком RD моделей

Все существующие на данный момент международные стандарты кодирования видео регламентируют синтаксис и семантику закодированного потока, но оставляют большой простор при выборе алгоритма создания этого потока, то есть алгоритмов кодирования. Выбор конкретного алгоритма позволяет кардинальным образом влиять как на качество кодирования, так и на необходимые для этого вычислительные ресурсы.

Центральное место среди алгоритмов управления видеокодеков занимают алгоритмы управления битрейтом кодека (rate-control algorithm). Основной задачей этих алгоритмов является максимизация заданного критерия качества кодирования при соблюдении ограничений на размер закодированных данных. Выделяют три основных уровня алгоритмов управления битрейтом [12]: уровень груп-

пы кадров, уровень кадров и уровень макроблоков. На уровне групп кадров (group of pictures, GOP) решаются задачи определения границ группы кадров и выбора для каждой группы целевого битрейта. Алгоритмы уровня кадра определяют тип отдельных кадров и их целевые размеры. На уровне макроблоков решаются задачи выбора параметра квантования и других параметров кодирования отдельных макроблоков.

В общем задачу всех этих алгоритмов можно формально поставить таким следующим способом:

$$Q_{opt} = \arg \min_{Q: R(Q) < R_{max}} D(Q),$$

где Q – параметры квантования объекта (кадра или макроблока), $R(Q)$ – размер объекта после кодирования, $D(Q)$ – внесённые искажения после квантования, R_{max} – целевой размер, Q_{opt} – искомые оптимальные параметры квантования.

Эта задача часто решается при помощи метода RDO (Rate Distortion Optimization) [1]. Метод сводится к выбору метода кодирования, минимизирующего следующее выражение (Лагранжиан):

$$J(Q) = R(Q) + \lambda D(Q),$$

где R – размер макроблока после кодирования, D – искажения после квантования (как правило, используется среднеквадратичное отклонение), λ – множитель Лагранжа, однозначно соответствующий целевому размеру объекта кодирования.

В этом выражении явным образом возникают две функции $R(Q)$ и $D(Q)$, называемые RD моделями. Эффективность многих предложенных на данный момент алгоритмов напрямую зависит от точности используемых RD моделей. Всегда есть возможность получить наилучшие данные при помощи имитации кодирования [9], но в ряде случаев такой подход невозможен из-за слишком большой вычислительной сложности. Альтернативой является задание аналитических моделей, приближающие RD характеристики реальных данных. Несмотря на потенциальные ошибки предсказания, аналитические модели намного более применимы в реально используемых видеокодеках.

Большинство RD моделей имеют дополнительные параметры. Наличие реальных RD характеристик объектов после их кодирования позволяет динамически подбирать дополнительные параметры модели в процессе кодирования. Таким образом, RD модели могут подстраиваться под реальные данные кодирования, повышая точность предсказания. Алгоритмы выбора адаптивных параметров являются неотделимой частью реализации RD моделей в видеокодеках. Взаимодействие RD моделей и кодека в общем виде показано на Рисунке 1.

1.2 Проблема оценки качества RD модели

В литературе на данный момент предложено большое количество различных RD моделей, однако их сравнение встречается крайне редко. Одной из возможных причин этого является сложность реализации нескольких моделей в одном кодеке и влияние на поведение RD модели множества деталей реализации кодека. Самым распространённым методом сравнения RD моделей на данный момент является наименьшее среднеквадратичное отклонение RD кривой предлагаемой RD модели с наилучшими параметрами от реальных данных кодирования. Этот метод используется в ряде статей, где производится сравнение предлагаемых или используемых RD моделей [6][7].

Ещё меньше внимания уделено вопросам анализа адаптивных параметров RD моделей и алгоритмам их обновления. Используются либо фиксированные параметры [7], либо параметр выбирается на основе предыдущего объекта кодирования [10], либо выбранный параметр минимизирует суммарную ошибку предсказания для нескольких предыдущих объектов [13].

В этой статье предлагаются новые методы оценки качества RD моделей. Выделены и рассмотрены три основных части RD моделей, анализ которых производится далее:

- Непосредственно RD функции (зависимости искажений и размера от параметра квантования). В большинстве литературы именно этому аспекту анализа RD моделей уделяется больше всего внимания.
- Алгоритмы обновления изменяемых параметров модели. Основная задача этих алгоритмов – адаптировать модель к реальным данным, которые могут быть получены после кодирования объектов кодеков.
- Алгоритм выбора реальных данных, которые могут быть использованы для анализа моделей. Основная задача – имитировать ту часть процесса кодирования кодеков, которая наиболее тесно связана с получением RD характеристик данных. При этом опускаются многие технические детали стандартов и реализации конкретных кодеков, что позволяет оценивать модель отдельно от использующего её кодека.

1.3 Предыдущие работы

Все используемые на данный момент алгоритмы управления кодеком с точки зрения моделирования RD характеристик можно разделить на две большие категории: использующие *эмпирические данные* и использующие *аналитические RD модели*.

Первые предварительно кодируют текущий объект (например, макроблок или кадр) с несколькими параметрами, а затем, в случае необходимости, строят аппроксимацию результатов для всех других параметров при помощи полученных данных. Так, в [8] минимизируется дисперсия искажений отдельных кадров при помощи аппроксимации RD функций. Для построения аппроксимации используются несколько реальных замеров искажений и битрейта для разных параметров квантования, значения между которыми интерполируются кусочно-непрерывными функциями. В ряде работ, учитывающих зависимости между объектами кодирования [9], авторы вычисляют RD функции непосредственно для каждого из возможных параметров кодирования. В целом эмпирические модели требуют значительных

вычислительных ресурсов и используются, когда точность предсказания критична.

В отличие от эмпирических алгоритмов, аналитические алгоритмы используют аналитические RD модели и могут получать оптимальные параметры кодирования в виде замкнутой формулы. Так, в [10], исходя из предположения о нормальном распределении коэффициентов межкадровой разницы после дискретного косинусного преобразования (ДКП), строится логарифмическая RD модель, используемая в дальнейшем для выбора квантайзеров при кодировании макроблоков внутри кадра. В тестовой модели MPEG “Test Model 5” [11] используется обратно квадратичная модель [13], являющаяся аппроксимацией логарифмической модели. Авторы [14] исходят из распределения Коши для коэффициентов ДКП. Получаемая при этом RD модель аппроксимируется экспоненциальными моделями, параметры которых могут быть вычислены из характеристик кадров видео. В [7] предлагается расширение логарифмической модели при помощи использования экспоненциального распределения коэффициентов.

Несмотря на многочисленные RD модели, предложенные на данный момент в литературе, вопросам выбора адаптивных параметров в RD моделях уделено намного меньше внимания. Так, в [10] авторы вводят в логарифмическую модель один адаптивный параметр, который вычисляется на основе обновления предыдущего значения в зависимости от результатов кодирования текущего кадра. В [13] при построении квадратичной RD модели используются два параметра, которые для текущего объекта кодирования подбираются таким образом, чтобы минимизировать ошибку предсказания на всех предыдущих макроблоках кадра. В [5] для вычисления адаптивных параметров используются параметры предыдущего кадра. В целом авторам неизвестны статьи, в которых сравниваются разные методы выбора адаптивных параметров моделей.

Остальная часть статьи организована следующим образом. В Разделе 2 описываются используемые при анализе RD модели, известные из литературы и используемые в современных кодексах, а так же предложенные авторами новые RD модели. В Разделе 3 приводятся примеры алгоритмов обновления параметров RD моделей. В Разделе 4 описываются примеры схем выбора реальных данных для адаптации моделей. Результаты проведённых авторами замеров качества RD моделей и краткие выводы из полученных результатов описываются в Разделе 5. Наконец, краткие итоги предложенных методов и проведённых исследований приводятся в завершающем Разделе 6.

2. ТЕСТИРУЕМЫЕ МОДЕЛИ

2.1 Используемые в кодексах RD модели

2.1.1 Классическая модель

Анализ RD моделей впервые проводился в теории кодирования сигналов с потерями. Общая постановка задачи может быть формализована следующим образом [2]:

$$R(D) = \min_{p(\hat{x}|x) \sum_{(x,\hat{x})} p(x) p(\hat{x},x) d(\hat{x},x) \leq D} I(X; \hat{X})$$

где $R(D)$ – искомая RD ф-ция, X – исходный сигнал, \hat{X} – декодированный сигнал, I – полная информация, p – плотность

вероятности сигнала, d – расстояние между исходным и декодированным сигналом.

Смысл этой формулы в том, что для сигнала ищется такой код, чтобы среднее искажение не превосходило заданной величины D , а на передачу тратилось минимальное количество ресурсов.

Можно показать, что если исходный сигнал имеет нормальное распределение $N(\mu, \sigma^2)$, RD функция имеет следующий вид:

$$R(D) = \sigma^2 2^{-2R},$$

Или, в терминах зависимости от линейного квантазера с порогом Δ :

$$D(\Delta) = \frac{\Delta^2}{\beta},$$

$$R(\Delta) = \frac{1}{2} \log_2 \left(\frac{\varepsilon^2 \beta \delta^2}{\Delta^2} \right) = C - \log(\Delta) + \frac{1}{2} \log(\delta^2),$$

где $\beta=12$, $\varepsilon=1.4$ для нормального распределения, $C = \frac{1}{2} \log(\varepsilon^2 \beta)$.

Несмотря на теоретическую обоснованность данной формулы, она является лишь теоретическим пределом при стремлении размера кодируемых единиц к бесконечности. Кроме того, реальные элементы смещённой межкадровой разности имеют скорее распределение Лапласа, чем нормальное [15]. Поэтому при кодировании реальных данных (отдельных кадров, макроблоков) эта модель может иметь довольно значительные ошибки.

2.1.2 Обратная модель

Зависимость размера R закодированного объекта от коэффициента квантования Q определяется следующим выражением:

$$R(Q) = \frac{C}{Q^\alpha},$$

где C – адаптивный параметр модели, α – константа.

В большинстве случаев полагают $\alpha=1$. Модель является довольно грубым приближением реальной зависимости, однако успешно используется на практике.

2.1.3 Линейная RD модель в ρ -домене

Основная идея RD моделей в ρ -домене – вместо функций $D(Q)$ и $R(Q)$ использовать функции $D(\rho)$ и $R(\rho)$, где $\rho = \rho(Q)$ может быть быстро вычислено. В своих работах [3][4] Zhihai He et al. предложили использовать в качестве домена количество ненулевых коэффициентов ДКП преобразования после квантования. Сложность вычисления зависимости $\rho(Q)$ весьма низкая – требуется всего лишь один проход по всем коэффициентам макроблока для формирования таблицы всех значений квантазера. Авторами предложены следующие формулы для RD модели:

$$R(\rho) = \Theta(1 - \rho)$$

и

$$D(\rho) = \sigma^2 e^{-\alpha(1-\rho)}.$$

где Θ и α – адаптивно подбираемые параметры.

В дальнейшем в работе [5] была предложена другая RD модель искажений в ρ -домене:

$$D(\rho) = \kappa \sigma^2 e^{-2\Theta(1-\rho)}$$

где κ – адаптивно подбираемый параметр.

Модели в ρ -домене показывают весьма значительную точность предсказания RD характеристик, однако обладают рядом недостатков, среди которых можно выделить следующие:

- для использования этих моделей необходимо произвести DCT преобразования макроблока;
- сложность вычисления $Q(\rho)$ выше, чем сложность вычисления прямой функции $\rho(Q)$.

Предложенные модели

Проводимый в этой статье анализ RD моделей позволяет не только анализировать уже используемые RD модели, но и предлагать новые решения. Далее рассматриваются две новые легко реализуемые в кодеках RD модели: табличная модель и модель в ρ -домене со смещением. Выводы об эффективности предложенных RD моделях можно сделать на основе результатов анализа в Разделе 5.

2.1.4 Табличная модель

Табличная модель предсказывает новые данные с использованием уже обработанных объектов кодирования. Создаётся сетка реальных значений функции $R(Q, \Sigma)$, где Q – квантазер, Σ – внутренние параметры макроблока. Новые данные предсказываются путём интерполяции уже хранящихся данных.

Авторами была реализована такая модель на основе двухмерной матрицы $R_{Q,s}$, где Q – квантазер, σ – дисперсия яркости блока. В каждой ячейке $R_{q,s}$ этой таблицы хранится среднее значение размера всех блоков, закодированных при использовании квантазера q , дисперсия которых удовлетворяла равенству

$$\frac{s}{M \cdot \sigma_{\max}} \leq \sigma < \frac{s+1}{M \cdot \sigma_{\max}},$$

где σ_{\max} – максимально рассматриваемая дисперсия пикселей блока, M – количество строк в матрице.

Предложенная модель не имеет адаптивных параметров, поэтому при анализе её результатов алгоритм обновления параметров для этой модели не указывается.

2.1.5 Модель в ρ -домене со смещением

Если рассмотреть зависимость $R(\rho)$ в стандартной модели в ρ -домене, то можно заметить, что при близких к единице значениях ρ размер блока падает тем быстрее, чем больше ρ .

Это происходит, поскольку при увеличении числа нулей в последовательности коэффициентов макроблока увеличивается и длина цепочки нулей в конце последовательности. Однако, эта цепочка нулей не кодируется, вместо этого кодеры вставляют в поток специальный код выхода, обозначающий тот факт, что все оставшиеся коэффициенты – нули. Таким образом, при увеличении количества нулей в последовательности коэффициентов длина завершающей цепочки нулей растёт все быстрее, что приводит к более

быстрому уменьшению размера с увеличением числа нулевых коэффициентов.

Этот эффект начинает играть существенную роль при значительном числе нулей среди всех коэффициентов макроблока. В этой ситуации возникает резкое снижение размера макроблока, и оставшаяся часть RD кривой, если ее продолжить не проходит через точку $(1,0)$.

Таким образом, для лучшей аппроксимации кривой $R(\rho)$ в модель может быть введена поправка, учитывающая вышеописанный эффект. Простейшим вариантом такой поправки является добавление дополнительного параметра в модель:

$$R(\rho) = \Theta(1 - \rho) + C$$

Примеры функции $R(\rho)$, приближающей ее классической модели в ρ -домене и модифицированной модели приведены на Рисунке 2.

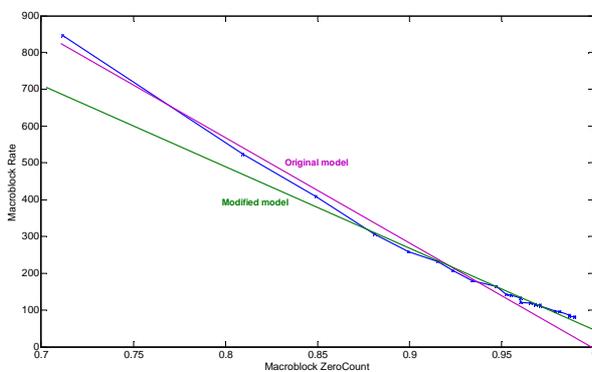


Рисунок 2: Модифицированная модель в ρ -домене.

3. АЛГОРИТМЫ ОБНОВЛЕНИЯ ПАРАМЕТРОВ

Адаптивные параметры используемых при кодировании RD моделей является мощным средством для повышения точности моделей. Однако в процессе кодирования часто нет возможности выбирать эти параметры оптимальным образом. Основная причина этого – слишком большая вычислительная сложность получения всех необходимых данных. Поэтому можно говорить об алгоритмах обновления параметров, которые являются неотъемлемой частью процесса использования моделей в кодеке.

Кроме того, доступные для адаптации модели данные также влияют на её точность. Например, в случае однопроходного кодирования, в кодеке для каждого макроблока или кадра доступны RD данные только для одного параметра квантования.

3.1 «По предыдущему»

Самый простой способ адаптивного подбора параметров. После окончания кодирования предыдущего объекта, для него становится известными внесённые искажения и требуемый размер при использованном квантаizerе. В случае однопараметрической модели этого достаточно, чтобы вычислить параметр модели. Именно этот параметр и предлагается использовать для предсказания искажений и размера текущего объекта.

В случае многопараметрических моделей этот алгоритм в явном виде не применим. Возможен вариант определения априорной зависимости между параметрами модели, но это, фактически, превращает многопараметрическую модель в однопараметрическую.

3.2 «По среднему»

В процессе кодирования становится известны результаты кодирования каждого из уже обработанных объектов. Алгоритм «по среднему» вычисляет параметр, минимизирующий отклонение модели на этом наборе объектов. Этот алгоритм хорош, если объекты кодирования имеют однородные RD характеристики.

3.3 «По соседям»

Эффективно оценить адаптивные параметры RD модели для текущего объекта можно по RD данным из объектов со схожими свойствами. Однако в естественных видеопоследовательностях соседние объекты (кадры, отдельные участки кадра, пиксели) очень похожи. Именно поэтому следует использовать последний закодированный объект, а не, скажем, первый из всей последовательности. Соседние с точки зрения кодирования объекты не всегда являются соседними с точки зрения исходной видеопоследовательности. Например, последовательно закодированные P-кадры могут быть разделены рядом B кадров, а последовательно закодированные макроблоки могут являться последним и первым макроблоком строк и иметь мало общего.

Таким образом, для вычисления параметров модели следует использовать объекты, являющиеся соседними в исходной последовательности, а не соседними в процессе кодирования.

В случае макроблоков, рассмотрим окрестность P всех закодированных блоков, расположенных в непосредственной близости от текущего макроблока. Для одного кадра в эту окрестность попадают блоки слева и сверху от текущего макроблока. В случае нескольких кадров могут быть использованы макроблоки с уже закодированных кадров с близкими пространственными координатами. Результирующей оценкой является среднее арифметическое параметров всех макроблоков, или наилучшие параметры для кодирования всех рассматриваемых объектов.

4. АЛГОРИТМЫ ВЫБОРА КОДИРУЕМЫХ ДАННЫХ

Итак, RD модели современных кодеков используют настраиваемые в процессе кодирования параметры, то есть являются адаптивными. Алгоритмы обновления параметров используют реальные данные для подбора параметров этих моделей. Однако остаётся открытым вопрос о том, какие именно данные подаются на вход алгоритму адаптивного изменения параметров. Далее в этом разделе рассмотрены различные способы выбора реальных данных. Эти способы фактически являются имитацией работы реальных алгоритмов управления кодеками, но абстрагированы от множества мелких деталей реализации стандартов кодирования видео и устройства самих видеокодеков.

4.1 Постоянный параметр квантования

Для адаптивного подбора параметров модели используются реальные данные только с одним фиксированным параметром квантования. Рассмотрим кодирование макроблоков внутри

одного кадра. Фиксируется квантайзер Q_{train} , с которым последовательно кодируются макроблоки. После кодирования очередного макроблока кодек получает информацию о реальной RD характеристике $R(Q_{train})$ только что закодированного макроблока. Эта информация передаётся в алгоритм подбора параметров RD модели.

Оценка качества модели в этом режиме проводится при помощи анализа ошибки при предсказании характеристик очередного макроблока для произвольного квантайзера Q_{test} . То есть, перед кодированием каждого макроблока замеряется ошибка RD модели для тестового квантайзера, после чего значения по всем макроблокам усредняются. Наиболее интересным является случай, когда $|Q_{train} - Q_{test}|$ не очень большой.

4.2 Постоянный битрейт

Задача кодирования с фиксированным битрейтом является важной для современных кодеков. Именно режим кодирования с жёсткими ограничениями на битрейт должен использоваться при передаче видео по спутниковым каналам или для трансляций в режиме реального времени.

Авторы использовали один из вариантов алгоритма удержания битрейта, при котором для кодировании каждого следующего макроблока оставшееся место для кадра разбивается равномерно между всеми ещё не закодированными макроблоками. То есть,

$$R_k = \frac{R - \sum_{i=0}^{k-1} R_i}{n - k},$$

где k – номер текущего макроблока, R_k – целевой битрейт для k -ого макроблока, n – количество макроблоков в кадре, R – целевой битрейт для кадра.

Для анализа ошибок модели используется наилучший возможный квантайзер при кодировании очередного макроблока (квантайзер, минимизирующий отклонение от целевого битрейта). Средняя ошибка относительно этого решения и является критерием качества модели в этом режиме кодирования.

4.3 Минимальная ошибка при ограниченном размере

Также, весьма важной задачей является кодирование последовательности кадров с минимальной суммарной ошибкой и ограничении на суммарный размер:

$$\min_{(q_1, \dots, q_n)} \sum_{i=1}^n D_i(q_i) \text{ при } \sum_{i=1}^n R_i(q_i) \leq R.$$

Эта задача является в точности задачей на нахождение условного минимума функции и обычно решается сведением к задаче поиска безусловного минимума функции путем использования множителей Лагранжа:

$$C = \min_{(q_1, \dots, q_n)} \left(\sum_{i=1}^n D_i(q_i) + \lambda \sum_{i=1}^n R_i(q_i) \right)$$

Здесь множитель Лагранжа λ задает соотношение между суммарной ошибкой кадров и их суммарным размером. Выбор множителя может производиться различными методами, рассмотрение которых выходит за рамки статьи. Будем предполагать, что параметр λ уже выбран к началу работы рассматриваемой части метода.

Многие алгоритмы управления кодеками действуют в предположении о независимости кодируемых макроблоков друг от друга, что позволяет перейти от минимизации общей суммы к минимизации стоимости на каждом макроблоке.

$$C = \sum_{i=1}^n \min_{q_i} (D_i(q_i) + \lambda R_i(q_i))$$

В этом случае алгоритм кодирования макроблоков предстает следующим образом. Сначала фиксируется параметр λ . Затем, для каждого макроблока находится квантайзер q_i , минимизирующий сумму $C_i = D_i(q_i) + \lambda R_i(q_i)$, где функции R_i и D_i есть предсказание моделью размера и ошибки на кодируемом макроблоке. После выбора квантайзера макроблок кодируется, и его реальные RD характеристики передаются в алгоритм обновления параметров модели.

5. ЭКСПЕРИМЕНТАЛЬНЫЕ РЕЗУЛЬТАТЫ

Таблица 1: Параметры видекодека для получения данных анализа RD моделей.

Часть кодека	Используемый алгоритм
Алгоритм компенсации движения	Полный перебор с пиксельной точностью, далее полупиксельное уточнение в радиусе 1 (ближайшие полупиксельные вектора)
Критерий минимизации ME	MSE (без учёта размера вектора)
Алгоритм выбора квантайзера	Константный квантайзер (как для кадров, так и для макроблоков)
Разбиение на слайсы	Каждая строка макроблоков – отдельный слайс
Структура типов кадров	Фиксированная: 1 I-кадр, 8 P-кадров

Для сбора исходных данных использовался видео кодек MPEG-2 из Intel IPP 5.0 Samples (примеры к библиотеке Integrated Performance Primitives от компании Интел) [16].

Стоит отметить, что благодаря использованию MSE в качестве критерия минимизации, а также полного пиксельного поиска, вектора для макроблоков внутри кадра выбирались независимо от векторов предыдущих макроблоков.

Для каждого макроблока в кодеке вычислялся размер и искажения для всех возможных квантайзеров. При расчёте размера не учитывалась стоимость кодирования векторов движения и заголовка макроблока (эта информация не учитывается в рассматриваемых RD моделях). Для вычисления размера производилось кодирование и декодирование макроблоков без записи результатов.

Таким образом, для каждой пары (макроблок, квантайзер) из кодека извлекалась следующая информация:

- размер макроблока после кодирования;
- искажения после декодирования (использовалось среднеквадратичное отклонение);

- количество нулевых коэффициентов ДКП после кодирования.

Далее эти данные анализировались средой анализа RD моделей, реализованной в виде набора скриптов для программного комплекса MATLAB. В этой среде были реализованы все описанные выше RD модели и методы обновления параметров.

Для получения данных использовался ряд последовательностей, состоящих как из стандартных последовательностей для тестирования видео (Foreman, Susi, Salesman), так и из сцен из видеофильмов (The Matrix, Terminator 2).

5.1 Постоянный квантайзер

Как уже было описано выше, в этом режиме кодирования имитировалось кодирование всех макроблоков с одинаковым квантайзером Q_b , результаты которого использовались для адаптации модели к реальным данным. Перед кодированием каждого из макроблоков сравнивались предсказания модели для каждого из квантайзеров и реальные данные. Эта информация сохранялась и усреднялась для всех рассмотренных макроблоков.

В итоге была получена информация о средней ошибке RD модели для кодирования с квантайзером Q_b при предсказании данных для квантайзера Q_p . Эти данные могут быть визуализированы как двумерная поверхность $E_{cap}(Q_b, Q_p)$. На Рисунке 3 и Рисунке 4 представлены примеры такой визуализации для ряда RD моделей на последовательности "Foreman" (независимо закодированный кадр).

Анализ других последовательностей позволяет сделать вывод, что в этом режиме кодирования можно выделить 3 RD модели с наилучшей ошибкой предсказания:

- Модифицированная модель в р-домене при больших коэффициентах квантования.
- Модель в р-домене при предсказании маленьких квантайзеров по данным малых квантайзеров
- Табличная модель при предсказании больших квантайзеров по данным малых квантайзеров

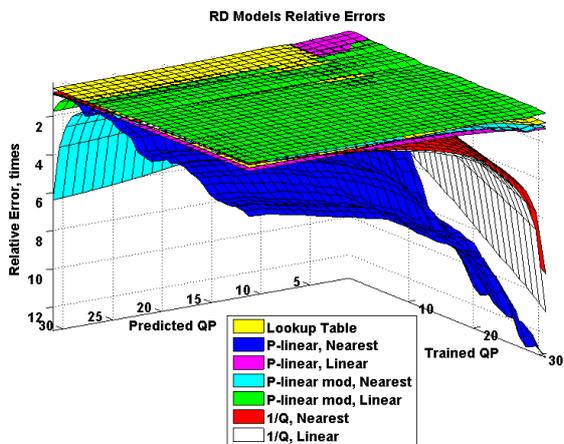


Рисунок 3: Точность всех RD моделей в режиме с постоянным квантайзером.

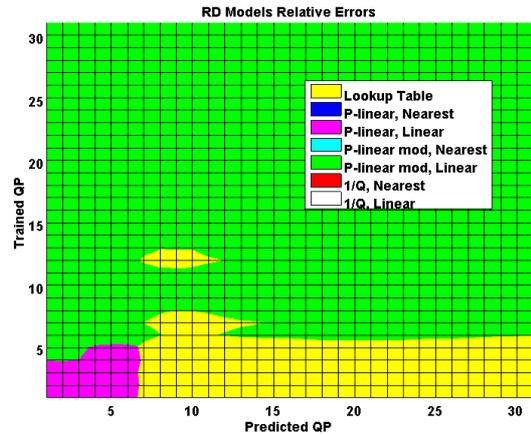


Рисунок 4: Точность всех RD моделей в режиме с постоянным квантайзером, лучшие модели (вид сверху).

Модифицированная модель в р-домене теряет стабильность при кодировании на малых квантайзерах, но обладает очень хорошей точностью при адаптации к сильно сжатым данным.

Использование алгоритма обновления параметров "по соседям" ухудшает поведение RD моделей. Возможной причиной данного факта может служить недостаточная стабильность этого алгоритма для работы с весьма однородными данными постоянного коэффициента квантования.

5.2 Постоянный битрейт

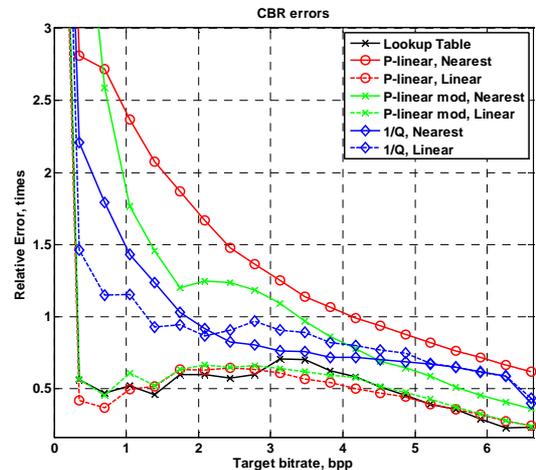


Рисунок 5: Все RD модели CBR I-кадры.

Основным критерием эффективности модели в этом режиме является функция $E_{cbr}(R_t)$ – зависимость средней ошибки модели от целевого битрейта R_t . При этом происходит усреднение ошибок по всем макроблокам кадра. Пример такой зависимости показан на Рисунке 5.

Анализ всех последовательностей позволяет сделать вывод о том, что среди рассмотренных моделей нет явных "лидеров" – примерно одинаковое качество показывают следующие RD модели:

- модифицированная RD модель в ρ -домене;
- RD модель в ρ -домене;
- табличная RD модель.

Как и в случае с постоянным квантайзером, в режиме кодирования с постоянным битрейтом алгоритм обновления параметров "по соседям" ухудшает поведение моделей.

Наибольшее ошибки предсказания имеет RD модель в ρ -домене при использовании алгоритма обновления "по соседям".

5.3 Максимальное качество при ограниченном размере

В качестве основного результата в этом режиме тестирования выступала функция $D(R)$ – зависимость среднего искажения кадра от битрейта. Примеры таких зависимостей показаны на Рисунках 6 и 7.

Анализ показывает, что результаты несколько различаются для различных типов кадров. Так, для I-кадров можно сделать следующие выводы:

- табличная модель имеет наименьшие ошибки предсказания;
- ошибки обратной модели незначительно выше, причём их величина практически не зависит от используемого алгоритма выбора параметров;
- наибольшие ошибки предсказания имеет модифицированная модель в ρ -домене.

Для P-кадров:

- наименьшие ошибки предсказания демонстрирует табличная модель и модель в ρ -домене с предсказанием по среднему;
- наибольшие ошибки предсказания имеет обратная модель с предсказанием по соседям.

На всех типах кадров модифицированная модель в ρ -домене с предсказанием по соседям оказывается очень неустойчивой и показывает однозначно плохие результаты.

5.4 Реализация моделей в видеокодеке

Для проверки адекватности предложенных тестов авторы реализовали некоторые из рассматриваемых RD моделей и алгоритмов обновления их параметров в видеокодеке Intel IPP MPEG-2 [16].

Для проверки работы были использованы те же настройки кодека, что и в случае с анализом RD моделей (см. Таблицу 1). В качестве тестовой видеопоследовательности была использована стандартная последовательность Foreman.

На Рисунке 8 показан пример RD характеристик рассматриваемого кодека с использованием протестированных моделей. Алгоритм управления кодека имитировал режим CBR, описанный в Разделе 4.2.

Результаты работы кодека подтверждают результаты, полученные с использованием предложенного метода анализа RD моделей. В частности, RD модель в ρ -домене даёт лучший результат в случае использования линейного предсказания, а характеристики RD модели для разных типов алгоритмов предсказания сближаются при увеличении битрейта. Другие

проведённые тесты подтверждают согласованность предложенных методов с реальными данными работы кодеков.

6. ЗАКЛЮЧЕНИЕ

В этой статье предложен новый метод оценки качества RD моделей, основанный на анализе ошибок предсказания характеристик закодированных данных этими моделями в определённых стандартных ситуациях. Использование достаточно общих алгоритмов анализа позволяет отказаться от учёта множества мелких деталей реализации кодеков, делая их более универсальными и подходящими для использования с кодеками почти всех современных стандартов видео.

Помимо непосредственно RD модели, проведён анализ алгоритмов обновления адаптивных параметров моделей. Практика использования RD моделей в видеокодеках показывает, что эти алгоритмы в значительной степени влияют на эффективность кодирования. Таким образом, при использовании моделей в реальных кодеках стоит думать так же о соответствующих алгоритмах их адаптации к реальным данным.

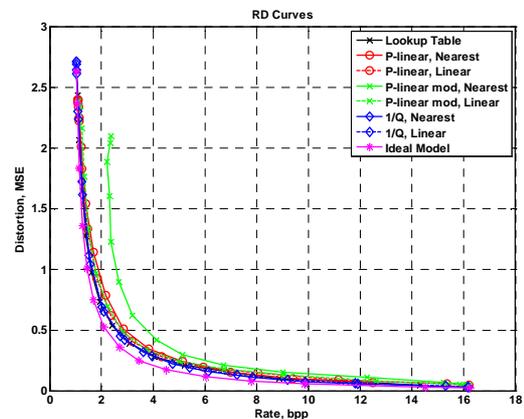


Рисунок 6: Все RD модели в режиме максимизации качества при ограничении на размер, ключевые кадры (I-кадры).

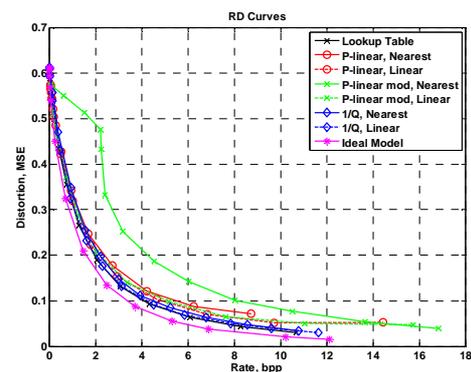


Рисунок 7: Все RD модели в режиме максимизации качества при ограничении на размер, зависимые кадры (P-кадры).

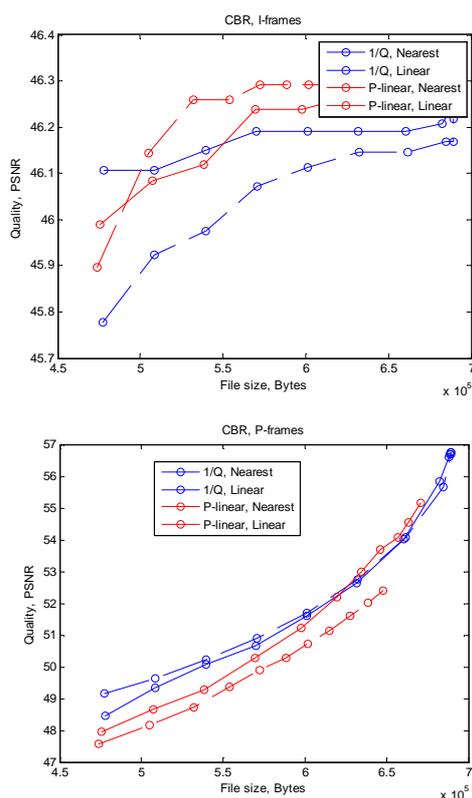


Рисунок 8: Реализация двух RD моделей и двух алгоритмов обновления параметров в кодеке Intel IPP MPEG-2 [16].

Проведённые исследования позволяют не только делать выводы относительно уже известных RD моделях, но и предлагать новые. Так, рассмотренные в статье табличная модель и модифицированная модель в p -домене могут быть использованы в определённых ситуациях для уменьшения ошибок предсказания. Ещё большее увеличение эффективности кодеков может быть получено при использовании гибридных RD моделей.

7. БЛАГОДАРНОСТИ

Авторы выражают благодарность Лаборатории Компьютерной Графики и Мультимедии факультета ВМиК МГУ имени Ломоносова и лично её заведующему Ю.М. Баяковскому. Так же авторы благодарят Videогруппу лаборатории и её руководителя Д.С. Ватолина за помощь в проведении исследований и множество советов при написании текста статьи.

8. СПИСОК ЛИТЕРАТУРЫ

- [1] G. J. Sullivan and T. Wiegand. *Rate-Distortion Optimization for Video Compression*. Nov. 1998, *IEEE Signal Processing Magazine*, pp. 74--90.
- [2] Thomas M. Cover, Joy A. Thomas. *Elements of information theory*. 2nd Edition. New York: Wiley-Interscience, 2006. ISBN 0-471-24195-4.
- [3] Yong Kwan Kim, Zhihai He, Mitra, S.K. *A novel linear source model and a unified rate control algorithm for H.263/MPEG-2/MPEG-4*. ICASSP 2001, pp. 1777--1780 vol.3.
- [4] Z. He and S. K. Mitra. *Optimum bit allocation and accurate rate control for video coding via p -domain source modeling*. *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 12, no. 10, Oct. 2002.
- [5] Kamaci N., Altunbasak Y. *P-domain rate-distortion optimal rate control for DCT-based video coders*. 2004, ICASSP, pp. iii-149-52 vol.3.
- [6] Zhihai He, Yong Kwan Kim, and Sanjit K. Mitra. *P-domain source modeling and rate control for video encoding and transmission*. ICASSP 2001, May 2001.
- [7] Min Dai, Dmitri Loguinov. *Analysis of Rate Distortion Functions and Congestion Control in Scalable Internet Video Streaming*. NOSSDAV 2003.
- [8] Liang-Jin Lin, Antonio Ortega. *Bit-Rate Control Using Piecewise Approximated Rate-Distortion Characteristics*. *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 8, No. 4, August 1998.
- [9] Ramchandran, K. Ortega, A. Vetterli, M. *Bit allocation for dependent quantization with applications to multiresolution and MPEG video coders*. *IEEE Transactions on Image Processing*, 1994.
- [10] Bo Tao, Bradley W. Dickinson, and Heidi A. Peterson, *Adaptive Model-Driven Bit Allocation for MPEG Video Coding*. *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 10, no. 1, February 2000.
- [11] *Coded Representation of Pictures and Audio Information, Test Model 5, ISO-IEC/JTC1/SC29/WG11*, 1993.
- [12] Nejat Kamaci, Yucel Altunbasak, Russell M. Mersereau. *Frame Bit Allocation for the H.264/AVC Video Coder Via Cauchy-Density-Based Rate and Distortion Models*. *IEEE Transactions On Circuits And Systems For Video Technology*, vol. 15, no. 8, August 2005.
- [13] Tihao Chiang; Ya-Qin Zhang. *A new rate control scheme using quadratic rate distortion model*. *Circuits and Systems for Video Technology, IEEE Transactions on*, vol.7, no.1, pp.246-250, Feb 1997.
- [14] Kamaci, N.; Altunbasak, Y.; Mersereau, R.M. *Frame bit allocation for the H.264/AVC video coder via Cauchy-density-based rate and distortion models*. *Circuits and Systems for Video Technology, IEEE Transactions on*, vol.15, no.8, pp. 994-1006, Aug. 2005.
- [15] Smoot S. R. and Rowe L. A. *Study of DCT coefficient distributions*. In *Proceedings of the SPIE Symposium on Electronic Imaging*, volume 2657, San Jose, CA, January 1996.
- [16] Библиотека программ Intel Integrated Performance Primitives. <http://www.intel.com/cd/software/products/asmona/eng/perflib/ipp/302910.htm>

Об авторах

Алексей Анатольевич Носков – студент факультета вычислительной математики и кибернетики Московского Государственного Университета им. М.В. Ломоносова
Адрес: Москва, 119899, Воробьевы горы, МГУ, 2-й учебный корпус, факультет ВМиК, кафедра автоматизации систем вычислительных комплексов.

E-mail: anoskov@graphics.cs.msu.ru

Александр Евгеньевич Паршин – аспирант факультета вычислительной математики и кибернетики Московского Государственного Университета им. М.В. Ломоносова
Адрес: Москва, 119899, Воробьевы горы, МГУ, 2-й учебный корпус, факультет ВМиК, кафедра автоматизации систем вычислительных комплексов.

E-mail: aparshin@graphics.cs.msu.ru

New analysis method of videocodec RD models and their adaptive paramethers

Abstract

Many videocodec control algorithms are based on models, which describe dependence between compression rate and quantity of introduced distortion (Rate-Distortion models). The issue of those models quality analysis independed on codec details and coding standart is raised in this paper. Additional to models itself, adaptive models parameters are also under consideration. Number of commonly used and proposed by authors Rate-Distortion models were analyzed, conclusions about models efficiency and possibilities of their improvements were made.

Keywords: *Rate-Distortion models, videocodec analysis, rate control algorithms.*

About the authors

Alexey Noskov is a student at Moscow State University, Department of Computational Mathematics and Cybernetics. His contact email is anoskov@graphics.cs.msu.ru.

Alexander Parshin is a PhD student at Moscow State University, Department of Computational Mathematics and Cybernetics. His contact email is aparshin@graphics.cs.msu.ru.