

УПРАВЛЕНИЕ БИТОВОЙ СКОРОСТЬЮ ГРУППЫ ИСТОЧНИКОВ ВИДЕОИНФОРМАЦИИ НА ОСНОВЕ СТАНДАРТА H.264/SVC

¹Беляев Е.А., ²Тюрликов А.М.

¹Санкт-Петербургский институт информатики и автоматизации РАН

²Санкт-Петербургский государственный университет аэрокосмического
приборостроения

В некоторых системах передачи видеоинформации обрабатывается группа из нескольких видеоисточников. Например, в системах цифрового телевизионного вещания, системах видеонаблюдения и др., осуществляется передача нескольких видеосигналов по общему беспроводному каналу связи. Так как всё большее распространение в таких системах получает расширение стандарта H.264/SVC [1] (Scalable Video Coding Extension of the H.264/AVC Standard) актуальна задача управления битовой скоростью группы видеоисточников на базе данного стандарта.

При традиционном подходе пропускная способность канала связи равномерно делится между видеоисточниками. Такой подход может привести к существенной разнице уровней искажения между обрабатываемыми видеоданными: одни видеопоследовательности будут передаваться с неоправданно высоким визуальным качеством, при этом другие видеопоследовательности будут иметь неприемлемый для просмотра уровень искажения.

Для устранения этого недостатка в настоящей работе предлагается динамически перераспределять ресурс канала связи в соответствии со статистическими свойствами различных источников видеоинформации. В отличие от других работ, перераспределение ресурса основано на минимаксном критерии искажения.

Доклад построен следующим образом. Сначала рассматривается случай управления битовой скоростью, когда кадры кодируются независимо друг от друга, без использования компенсации движения (режим Intra). Для этого случая формулируется оптимизационная задача и указывается способ управления битовой скоростью, который является решением этой оптимизационной задачи. Затем рассматривается случай зависимого кодирования кадров (режим Inter). При этом используется видоизменение оптимизационной задачи, с учетом особенностей расширения стандарта H.264/SVC.

Пусть кадры кодируются независимо друг от друга, без использования компенсации движения. В этом случае оптимизационную задачу управления битовой скоростью можно сформулировать следующим образом. Имеется S видеоисточников. Обозначим через $d(s, i)$ и $r(s, i)$ уровень искажения и битовые затраты на кадр с номером i видеоисточника с номером s . Традиционно, соотношение между уровнем искажения и битовыми затратами на кадр связывают следующим образом [2]:

$$d(s, i) = \sigma^2(s, i) \cdot \exp\left(-\frac{r(s, i)}{\beta(s, i)}\right), \quad (1)$$

где величина $\sigma^2(s, i)$ может быть интерпретирована как уровень искажения в худшем случае, если кадр i не передается. Величину $\beta(s, i)$ называют «сложностью» видеоисточника. Чем больше $\beta(s, i)$, тем большее количество бит требуется для обеспечения заданного уровня искажения. Последовательность кадров в каждой видеопоследовательности поделим на группы кадров. Обозначим количество кадров в каждой группе через M . В этом случае группа кадров с номером j состоит из

номеров кадров $G_j = \{j \cdot M, \dots, (j+1) \cdot M - 1\}$. Тогда для каждого кадра i входящего в группу кадров G_j необходимо выбрать такие параметры кодирования [3], чтобы

$$\begin{cases} \text{минимизировать } \max_{s \in S, i \in G_j} d(s, i) \\ \sum_{s \in S} \sum_{i \in G_j} r(s, i) = \sum_{s \in S} \sum_{i \in G_j} -\beta(s, i) \cdot \ln \left(\frac{d(s, i)}{\sigma^2(s, i)} \right) \leq R_{GOP} = \frac{M}{f} \cdot c, \end{cases} \quad (2)$$

где f - кадровая скорость, c - скорость канала связи, R_{GOP} - максимальное количество бит, которое может быть использовано для группы кадров G_j всех источников.

Можно показать, что $d(s, i) = d^*$ является решением задачи, где величина d^* вычисляется из выражения

$$\sum_{s \in S} \sum_{i \in G_j} -\beta(s, i) \cdot \ln \left(\frac{d^*}{\sigma^2(s, i)} \right) = R_{GOP}, \quad (3)$$

и равна

$$d^* = \exp \left(\frac{\sum_{s \in S} \sum_{i \in G_j} \beta(s, i) \cdot \ln \sigma^2(s, i) - R_{GOP}}{\sum_{s \in S} \sum_{i \in G_j} \beta(s, i)} \right). \quad (4)$$

Другими словами, решением задачи (2) является кодирование каждого кадра всех видеопоследовательностей с постоянным уровнем искажения d^* , которое является минимальным для заданной пропускной способности канала связи c . Справедливость данного утверждения следует из следующих рассуждений.

Предположим, что есть другие параметры кодирования при которых для всех кадров i и всех видеисточников s выполняется $d'(s, i) < d^*$. Из (1) следует, что $r'(s, i) > r^*(s, i)$, где $r^*(s, i)$ - битовые затраты на кадр i источника s при кодировании с искажением d^* . Тогда с учетом (3) должно выполняться следующее неравенство:

$$\sum_{s \in S} \sum_{i \in G_j} r'(s, i) > \sum_{s \in S} \sum_{i \in G_j} r^*(s, i) = R_{GOP},$$

из которого следует, что не существует других параметров кодирования, которые обеспечивают максимальный уровень искажения на кадр меньше чем d^* .

Теперь перейдем к случаю сжатия видеопоследовательностей, с использованием компенсации движения, при которой для каждого блока в текущем кадре выполняется поиск наиболее «похожего» блока в базовом кадре.

Для формирования масштабируемого потока видеоданных по кадровой скорости (temporal scalability) в расширении стандарта H.264/SVC применяется метод иерархически кодируемых В-кадров. На рис.1 показан пример для восьми кадров в группе. При таком подходе сначала кодируется кадр с номером 8, для которого базовым является кадр 0 из предыдущей группы. Затем кодируется кадр 4, для которого базовым является кадр 8, затем кадр 2, для которого базовыми являются кадры 0 и 4 и так далее. В результате образуются четыре уровня иерархии, каждый из которых соответствует своей кадровой скорости. Если $f = 30$, то уровню $T = 0$ соответствует кадр 8 и кадровая скорость 3.75, уровню $T = 1$ соответствует кадр 4 и кадровая скорость 7.5, уровню $T = 2$ соответствуют кадры 2 и 4 и кадровая скорость 15, уровню $T = 3$ соответствуют кадры 1, 3, 5, 7 и кадровая скорость 30.

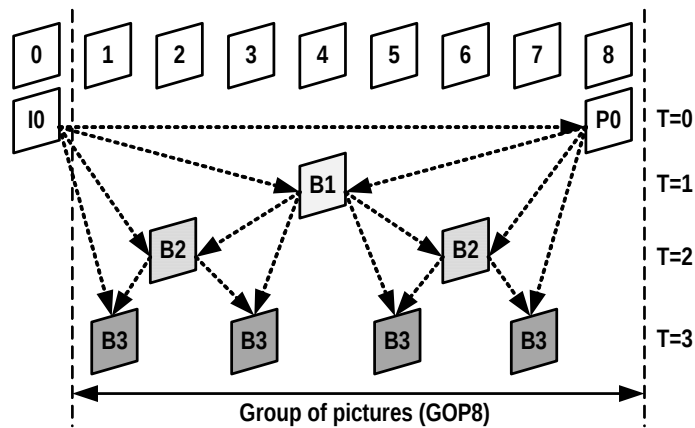


Рис.1 Схема иерархически кодируемых В-кадров в стандарте H.264/SVC

Из задачи зависимого квантования [4] следует, что при использовании компенсации движения для оптимального, с точки зрения кривой скорость/искажение, кодирования шаг квантования для базового кадра должен быть меньше шага квантования, используемого при кодировании текущего кадра. При этом уровень искажения базового кадра оказывается меньше, чем для компенсируемого кадра. Другими словами, для кодеров, использующих компенсацию движения, кодирование с постоянным уровнем искажения на кадр (3) не является наилучшим.

В случае расширения стандарта H.264/SVC, для примера на рис.1, наибольший уровень искажения возникнет у кадров, соответствующих уровню иерархии по кадровой скорости $T = 3$. Следовательно, в задаче (2) достаточно минимизировать наибольшее искажение только для кадров, соответствующих уровню иерархии $T = 3$. По аналогии с (4) будем исходить из предположения, что кодирование кадров на уровне иерархии $T = 3$ всех видеопоследовательностей с постоянным и наименьшим уровнем искажения, при котором сумма битовых затрат на все кадры в группе не превысит R_{GOP} бит будет решением оптимизационной задачи (2).

Кроме масштабирования по кадровой скорости расширение стандарта H.264/SVC предусматривает масштабирование по искажению (SNR scalability). При таком подходе первый уровень иерархии $l = 1$ формируется при помощи базового профиля H.264/AVC. При кодировании уровня $l = 2$ используются восстановленные кадры из уровня $l = 1$ (coarse grain scalability mode). При этом шаг квантования для уровня $l = 2$ меньше чем для уровня $l = 1$.

Для того, чтобы расширить оптимизационную задачу (2) для данного случая в настоящей работе предлагается для каждого уровня иерархии по искажению с номером l выделить R_{GOP}^l бит, причем

$$R_{GOP}^1 + R_{GOP}^2 + \dots + R_{GOP}^L = R_{GOP},$$

где L - число уровней иерархии по искажению, $R_{GOP}^l = \alpha_l \cdot R_{GOP}$, $\sum_{l=1}^L \alpha_l = 1$. В

результате решения задачи (2) для каждого уровня иерархии по искажению, максимальное искажение на кадр для уровня иерархии l каждого источника видеoinформации будет одинаковым.

Для получения практических результатов был использован открытый JSVM кодек [5], версии 9.15, который поддерживает расширение стандарта H.264/SVC. Результаты получены для первых 60-ти кадров известных тестовых видеопоследовательностей «Akiyo», «Hall» и «Football», разрешением 352x288, $f = 30$. В качестве меры искажения $d(s, i)$ выбрано пиковое отношение сигнал/шум

(PSNR). На рис.2 показаны значения PSNR для случая, когда на каждый видеисточник выделяется одинаковая доля канала связи равная 1000 кбит/с. При этом $\alpha_1 = 0.2$, $\alpha_2 = 0.8$, что означает, что на уровень иерархии $l=1$ выделено 200 кбит/с и на уровень иерархии $l=2$ выделено 800 кбит/с. Из рис.2 видно, что оба уровня иерархии по искажению ($l=1, l=2$) для видеопоследовательности «Akiyo» обеспечивают очень низкий уровень искажения ($PSNR > 44$ дБ), при этом уровень искажения для видеопоследовательности «Football» оказывается неприемлемо высоким для обоих уровней ($PSNR < 30$ дБ).

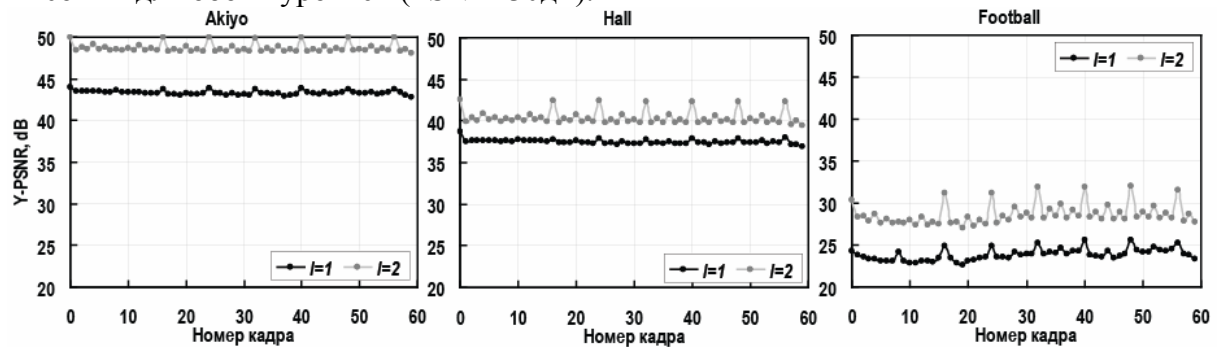


Рис.2 Одинаковая доля канала на каждый видеисточник

На рис. 3 показаны результаты использования предложенного метода управления для группы источников видеoinформации в случае, когда на три источника выделяется 3000 кбит/с. Из рис.3 следует, что данный подход позволяет обеспечить одинаковое минимальное значение PSNR для каждого уровня. При этом, за счет предложенного метода перераспределения ресурсов канала связи, обеспечивается комфортный для просмотра уровень искажения для всех видеопоследовательностей.

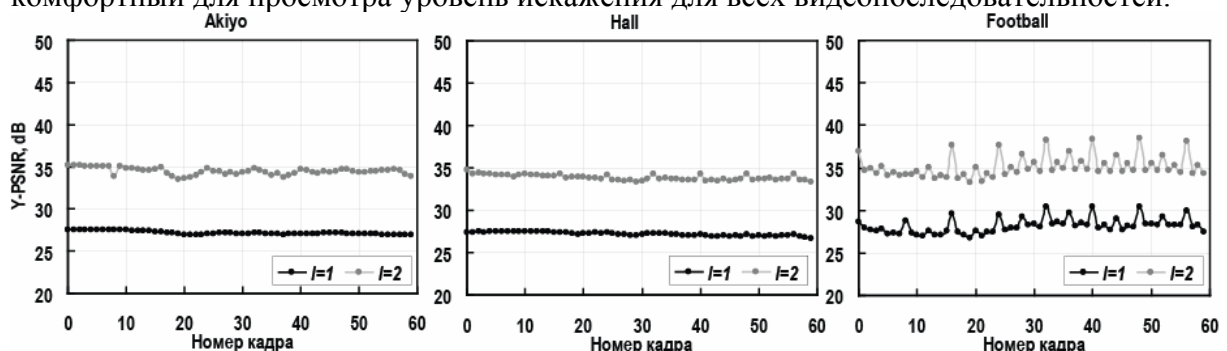


Рис.3 Распределение ресурса канала по минимаксному критерию искажения

Работа выполнена при поддержке Фонда содействия развитию малых форм предприятий в научно-технической сфере.

Литература

1. Advanced video coding for generic audiovisual services. *ITU-T Recommendation H.264 and ISO/IEC 14496-10 (AVC)*, 2009.
2. Dai M., Loguinov D., Radha H. Rate-distortion modeling of scalable video coders // *International Conference on Image Processing*, P.1093–1096, V.2, 2004.
3. E. Belyaev, T. Koski, J. Paavola, A. Turlikov and A. Ukhanova. Adaptive power saving on the receiver side in digital video broadcasting systems based on progressive video codecs // *The 11th International Symposium on Wireless Personal Multimedia Communications*, 2008.
4. H. Schwarz, D. Marpe, and T. Wiegand, Overview of the Scalable Video Coding Extension of the H.264 / AVC Standard // *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 17, No. 9, pp. 1103-1120, 2007.
5. JSVM 9.15 Software Package, CVS server for the JSVM software.

MULTI-SOURCE RATE CONTROL FOR VIDEO DATA BASED ON SCALABLE EXTENSION OF THE H.264/SVC STANDARD

¹Belyaev E.A., ²Turlikov A.M.

¹SPIIRAS, Russian Academy of Sciences

²Saint-Petersburg State University of Aerospace Instrumentation

This paper discusses transmission of several video streams in digital video broadcasting systems. Multi-source rate control based on MINMAX quality criteria is proposed. Results of practical application of the proposed algorithm in H.264/SVC video compression standard are presented.