
ОБРАБОТКА ВИДЕОИНФОРМАЦИИ

УДК 004.627

А. И. ВЕСЕЛОВ, М. Р. ГИЛЬМУТДИНОВ, Б. С. ФИЛИППОВ

МЕТОД ГЕНЕРАЦИИ СТОРОННЕЙ ИНФОРМАЦИИ ДЛЯ СИСТЕМ РАСПРЕДЕЛЕННОГО КОДИРОВАНИЯ ВИДЕОИСТОЧНИКОВ

На примере эталонной реализации кодека Discover описаны основные методы распределенного кодирования видеоисточников, проанализированы их преимущества и недостатки. Предложен новый метод генерации сторонней информации, основанный на иерархической оценке движения. Продемонстрирована эффективность кодека, использующего предложенный метод.

Ключевые слова: сжатие видеофайлов, распределенное кодирование, временная (межкадровая) интерполяция видеоданных, оценка движения.

Введение. Распределенное кодирование видеоданных (Distributed Video Coding, DVC) является сравнительно новым подходом к сжатию видеоисточника. Несмотря на то что теоретические предпосылки DVC были получены еще в 1970-е гг. [1, 2], активные исследования в данном направлении начались только в конце 1990-х гг. Это связано в основном с тем, что до недавнего времени распределенное кодирование рассматривалось как задача, не имеющая реального практического применения. Развитие технологий мобильной передачи данных привело к формированию новых требований к кодеру, учитывающих особенности мобильных видеоисточников: ограниченный, зачастую трудновосполняемый, объем аккумуляторной батареи и сравнительно малые вычислительные возможности мобильных устройств. Эти особенности накладывают существенные ограничения на допустимую сложность процедуры сжатия видеоданных на стороне мобильного передатчика. Кроме того, следует отметить, что описанные ограничения не принимались в расчет при разработке современных стандартов сжатия видеоданных ITU-T H.26x и ISO/IEC MPEG [3]. В связи с этим разработка подходов к сжатию видеоданных, учитывающих описанные выше особенности, является актуальной задачей.

Для решения этой задачи во многих работах в качестве перспективной технологии указывается DVC, поскольку в ее архитектуре заложено снижение сложности кодирования при сохранении степени сжатия. Основная сложность при этом переносится со стороны кодера на декодер: наиболее вычислительно сложная операция устранения временной избыточности, связанная с формированием ошибок предсказания на стороне кодера, заменяется процедурой интерполяции (или экстраполяции) кадров на стороне декодера. Блок, отвечающий за предсказание кадров на стороне декодера, принято называть блоком *генерации сторонней информации*. Ошибки, которые могут быть внесены в кадры в процессе интерполяции, исправляются с помощью проверочных битов некоторого помехоустойчивого кода, досылаемых кодером по запросам от декодера. Степень сжатия определяется количеством проверочных битов.

Из сказанного выше следует, что на эффективность устранения временной избыточности на стороне декодера и в конечном итоге — на сжатие будут влиять следующие факторы:

- точность метода генерации сторонней информации;
- эффективность исправления ошибок предсказания.

В настоящей статье рассматривается задача генерации сторонней информации, проанализирован метод генерации сторонней информации, используемый в эталонной реализации DVC кодека Discover [4], предложен подход, использующий более точную процедуру оценки движения.

Метод генерации сторонней информации, используемый в кодеке Discover. Наиболее популярным методом генерации сторонней информации является временная интерполяция кадров. В основе алгоритмов интерполяции лежат процедуры оценки и компенсации движения, позволяющие определить координаты одинаковых объектов на базовых кадрах и рассчитать положение этих объектов на промежуточном кадре. В кодеке Discover используется блоковая оценка движения [5], реализуемая в два этапа (рис. 1).

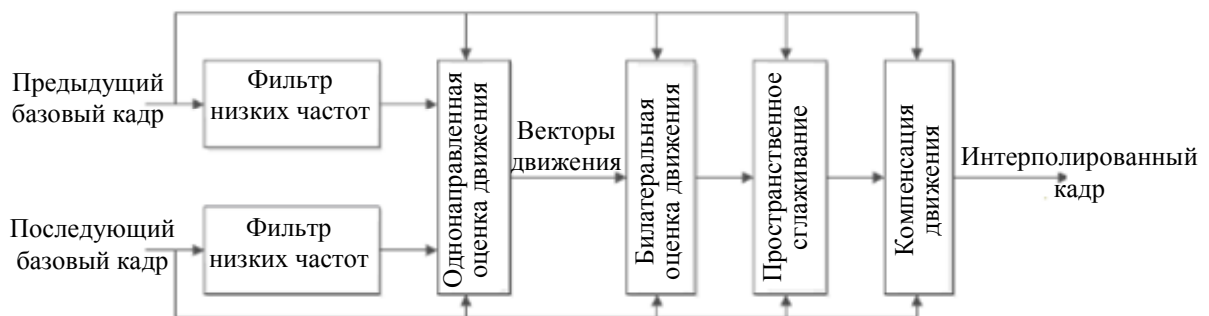


Рис. 1

Перед оцениванием движения базовые кадры подвергаются низкочастотной фильтрации с целью подавления шумов и увеличения точности последующего поиска. После однонаправленной оценки выполняется уточняющая *билатеральная* [6, 7] оценка с меньшим радиусом поиска, по завершении которой каждой координате на интерполированном кадре будет поставлен в соответствие вектор движения, т.е. на интерполированном кадре не будет участков с „коллизиями“ и „дырами“, которые возможны после первого шага [6]. Следующим шагом алгоритма является пространственное сглаживание множества векторов, объединенных в так называемое *векторное поле*. Для этого используется взвешенная медианная фильтрация. Для блока с координатами $\mathbf{p}_i = \begin{pmatrix} p^x \\ p^y \end{pmatrix}$ (под координатами блока

понимаются координаты верхнего левого угла блока по вертикали p^y и по горизонтали p^x)

и вектором $\mathbf{mv}_i = \begin{pmatrix} mv^x \\ mv^y \end{pmatrix}$ эта операция определена как:

$$\mathbf{mv}'_i = \arg \max_{\mathbf{mv} \in \{\mathbf{mv}_1, \dots, \mathbf{mv}_N\}} \left[w_j \left(\sum_{j=1}^N \left| \|\mathbf{mv}_i - \mathbf{mv}_j\|_L - \|\mathbf{mv} - \mathbf{mv}_j\|_L \right| \right) \right],$$

где $\|\cdot\|_L$ — норма вектора, \mathbf{mv}'_i — результат фильтрации, $\{\mathbf{mv}_1, \dots, \mathbf{mv}_N\}$ — множество „связанных“ с \mathbf{mv}_i векторов, в это множество входят векторы соседних блоков, а также вектор для блока с координатой \mathbf{p}_i в предыдущем интерполированном кадре; w_j — весовые коэффициенты, определяемые как:

$$w_j = \frac{MSE(\mathbf{mv}_i, \mathbf{p}_i)}{MSE(\mathbf{mv}_j, \mathbf{p}_i)},$$

где $MSE(\mathbf{mv}, \mathbf{p})$ — оператор расчета среднего квадратического отклонения при применении вектора \mathbf{mv} к блоку с координатой \mathbf{p} .

Полученное сглаженное билатеральное векторное поле используется на последнем шаге алгоритма для компенсации движения и формирования интерполированного кадра:

$$\tilde{F}^t(\mathbf{p}_i) = \frac{1}{2} \left[F^{t-1}(\mathbf{p}_i - \alpha \mathbf{mv}) + F^{t+1}(\mathbf{p}_i + (1 - \alpha) \mathbf{mv}) \right],$$

где через $\tilde{F}^t(\mathbf{p}_i)$ обозначена интенсивность пиксела, находящегося на позиции \mathbf{p}_i в интерполированном кадре \tilde{F}^t .

Метод генерации сторонней информации, основанный на иерархической оценке движения. Одним из существенных недостатков применяемого в Discover метода генерации сторонней информации является то, что в нем используется обычная блоковая оценка движения с полным перебором всех возможных векторов движения в некотором множестве. Учет корреляции векторов соседних блоков осуществляется только при медианной фильтрации векторного поля, но не используется в самой процедуре оценки движения. Такой подход в большинстве случаев неудобен для оценки так называемого *истинного движения (True Motion)* объектов [8], которое является определяющим фактором при интерполяции кадров.

Рассмотрим предлагаемый метод генерации сторонней информации, основанный на процедуре иерархической оценки движения. Эта процедура показывает хорошие результаты при решении схожей задачи — интерполяции кадров в алгоритмах преобразования кадровой скорости [9, 10]. Метод основан на постепенном уточнении векторов движения (начиная от оценки глобального смещения в рамках всего кадра и заканчивая оценкой локальных смещений объектов) посредством *иерархической* обработки кадров, использующей совокупность кадров различного разрешения [11]. Альтернативным способом является изменение размеров блоков при оценке движения с сохранением разрешения кадров. Такой подход учитывает корреляцию векторов и позволяет более точно оценивать истинное движение объектов, чем реализованный в Discover метод.

При оценке истинного движения также используется многоэтапный поиск: на начальном этапе осуществляется обычная блоковая билатеральная оценка движения, затем полученная оценка уточняется за счет использования векторов смежных блоков [8]. Поскольку при оценке движения реальный объект в кадре разбивается на несколько блоков, использование связей между блоками позволяет исправить ошибки поиска. Перед проведением дополнительного поиска выполняется оценка *надежности векторов движения*, полученных в ходе начального поиска. Надежность вектора — это количественный критерий, характеризующий степень влияния данного вектора на разброс значений векторов в его области [12]. Дополнительное разбиение каждого уровня иерархической оценки позволяет повысить точность оценки истинного движения. В работе [13] показано, что поиск с использованием полного перебора всех возможных векторов движения для блока приводит к появлению ошибок в определении истинного движения. В связи с этим в предложенном методе генерации сторонней информации предлагается использовать подоптимальный поиск, например, *градиентный спуск* [14]. Также предлагаемый алгоритм использует предобработку базовых кадров для выделения регионов со статичными объектами [15]. Эта информация позволяет существенно повысить точность оценки движения. В качестве модуля компенсации движения используется компенсация движения *с перекрытиями* (Overlapped Block Motion Compensation, ОБМС), которая позволяет уменьшить блоковые искажения на интерполированном кадре [16]:

$$\tilde{F}^t(\mathbf{p}_i) = \frac{1}{2} \sum_{n=1}^5 W_n \otimes \left[F^{t-1}(\mathbf{p}_i - \alpha \mathbf{mv}_n) + F^{t+1}(\mathbf{p}_i + (1 - \alpha) \mathbf{mv}_n) \right],$$

где $\mathbf{mv}_1, \mathbf{mv}_2, \dots, \mathbf{mv}_5$ — вектор текущего блока \mathbf{p}_i и его четырех соседей, W_1, W_2, \dots, W_5 — матрицы весовых коэффициентов, размер матриц совпадает с размерами блоков

$$\sum_{i=1}^5 W_i = J,$$

где J — матрица, состоящая из всех единиц.

Окончательная схема предлагаемого метода генерации сторонней информации приведена на рис. 2.

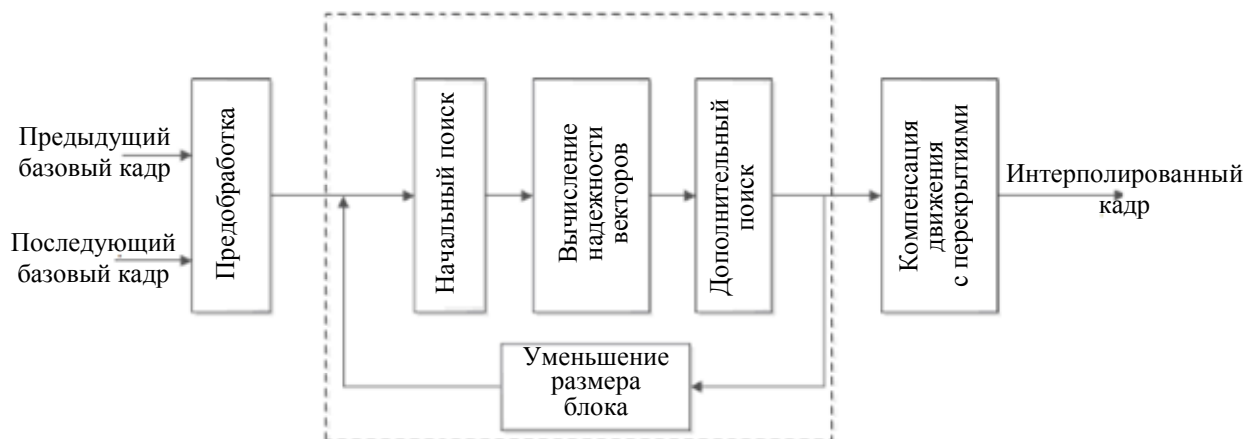


Рис. 2

Для оценки предлагаемого метода генерации сторонней информации был реализован видеокодек, работающий по следующей схеме [17]. Ключевые кадры обрабатываются, как в Discover. По восстановленным ключевым кадрам с помощью предложенного алгоритма генерируется сторонняя информация. Помехоустойчивое кодирование для исправления ошибок в сторонней информации не используется, т.е. промежуточные кадры обрабатываются только на стороне декодера и дополнительные биты от кодера не требуются. Интерполированный кадр выдается в выходной поток в качестве восстановленного кадра. Для оценки предложенной процедуры генерации сторонней информации были построены зависимости „скорость—искажение“ для различных кодеков на последовательностях из стандартного набора тестовых видеоданных [18]. В качестве критерия искажения использовалось усредненное по всем кадрам пиковое отношение сигнал/шум для яркостной компоненты (PSNR, Peak Signal-to-Noise Ratio). Скорость (Bitrate) рассчитывалась как среднее число килобитов, передаваемых в секунду. Частота следования кадров для тестовых видеопоследовательностей составляла 15 кадров в секунду, формат QCIF. В качестве сравниваемых кодеков были использованы (рис. 3):

- кодек стандарта H.264, режим Intra;
- кодек стандарта H.264, режим Inter;
- кодек стандарта H.264, режим Inter No Motion;
- кодек Discover;
- реализованный кодек (SUAI DVC).

Кодек H.264 Intra обрабатывает все кадры как ключевые. Остальные кодеки используют режим IBIB, т.е. каждый второй кадр обрабатывался как ключевой. В H.264 Inter предсказание осуществляется на кодере. Для H.264 Inter No Motion в качестве ошибки предсказания используется разница между базовым кадром и кодируемым.

На всех последовательностях, кроме Soccer, которая имеет низкую межкадровую корреляцию, SUAI DVC выигрывает у Discover в среднем на 0,5—1 дБ. При этом оба кодека проигрывают на всех последовательностях кодеку H.264 Inter (со сложным кодером). Если сравнивать кодеры низкой сложности, то в целом производительность DVC кодеков

приблизительно на том же уровне, что у H.264 Inter No Motion и, как правило, выше, чем у H.264 Intra.

В текущей версии кодека исправление ошибок предсказания не реализовано и, как следствие, дополнительные биты со стороны кодера для неключевых кадров не передаются. Это объясняет отсутствие по кодеку SUAI DVC для больших значений Y-PSNR.

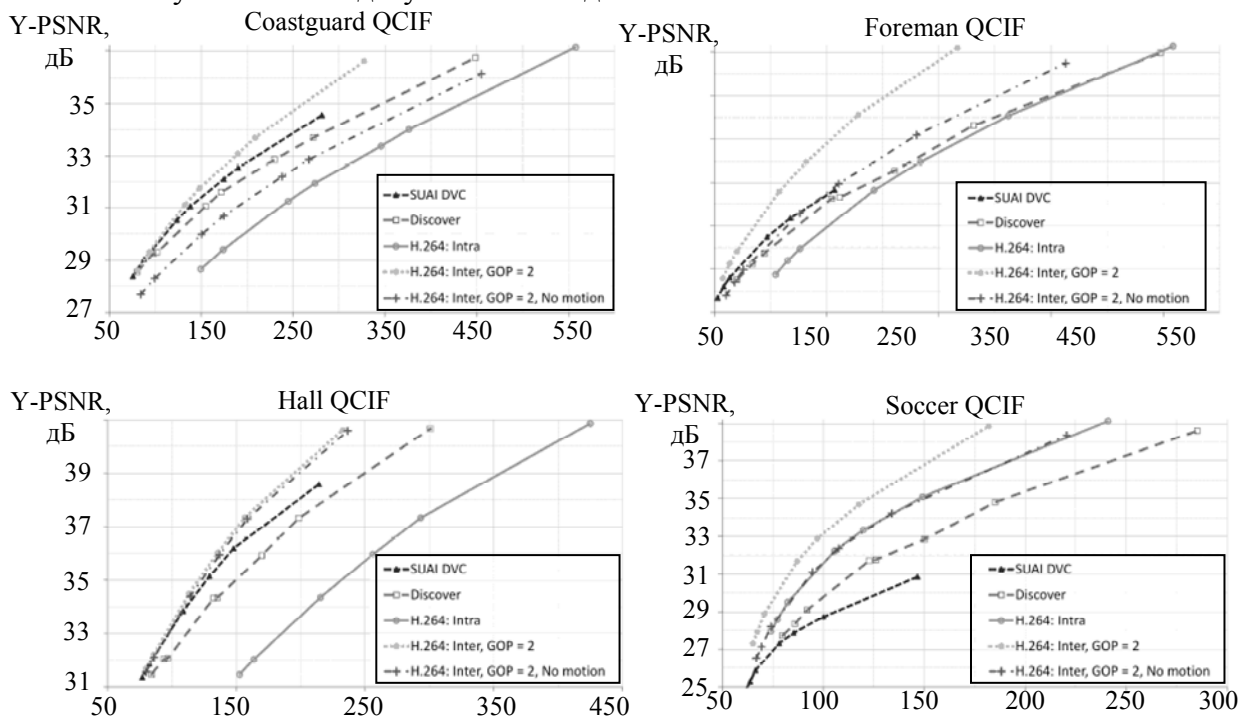


Рис. 3

Заключение. В работе приведено описание основных методов, используемых при распределенном кодировании видеоданных. Выделены наиболее значимые для этого подхода задачи и описаны способы их решения в эталонной реализации кодека Discover. Предложен новый более сложный метод генерации сторонней информации на стороне декодера, проведена его предварительная оценка. Реализован новый кодек, использующий предложенный метод. Показано, что для тестовых видеопоследовательностей с высокой межкадровой корреляцией реализованный кодек превосходит Discover в среднем на 0,5—1 дБ по критерию „скорость—искажение“. Таким образом, использование предложенного метода генерации сторонней информации позволяет повысить производительность DVC кодека без увеличения сложности кодирования.

СПИСОК ЛИТЕРАТУРЫ

1. Wolf J. K., Slepian D. Noiseless Coding of Correlated Information Sources // IEEE Transact. on Information Theory. 1973. Vol. 19, N 4. P. 471—480.
2. Ziv J., Wyner A. D. The Rate-Distortion Function for Source Coding with Side Information at the Decoder // IEEE Transact. on Information Theory. 1976. Vol. 22. P. 1—10.
3. Advanced video coding for generic audiovisual services. Recommendation ITU-T H.264 ISO/IEC 14496-10. 2003. 732 p.
4. DISCOVER project page [Электронный ресурс]: <www.discoverdvc.org>.
5. Pereira F., Ascenso J., Brites C. Improving Frame Interpolation with Spatial Motion Smoothing for Pixel Domain Distributed Video Coding // Proc. of 5th EURASIP Conf. on Speech and Image Processing, Multimedia Communications and Services. 2005. P. 252—257.

6. Choi B.-T., Lee S.-H., Ko S.-J. New Frame Rate Up-Conversion Using Bi-Directional Motion Estimation // IEEE Transact. on Consumer Electronics. 2000. Vol. 46. P. 603—609.
7. Марковский С. Г., Тюрликов А. М. Использование идентификаторов абонентов для резервирования канала множественного доступа // ИУС. 2008. № 2(33). С. 28—35.
8. de Haan G. et al. True Motion Estimation with 3-D Recursive Search Block-Matching // IEEE Transact. on Circuits and Systems for Video Technology. 1993. Vol. 3. P. 368—379.
9. Lee Y.-L., Nguyen T. Method and Architecture Design for Motion Compensated Frame Interpolation in High-Definition Video Processing // IEEE Intern. Symp. on Circuits and Systems (ISCAS). 2009. P. 1633—1636.
10. Jeon B.-W., Lee G.-I., Lee S.-H., Park R.-H. Coarse-to-fine Frame Interpolation for Frame Rate Up-Conversion Using Pyramid Structure // IEEE Transact. on Consumer Electronics. 2003. Vol. 49, N 3. P. 499—508.
11. Houlding D., Vaisey J. Pyramid Decompositions and Hierarchical Motion Compensation // Digital Video Compression: Algorithms and Technologies. Proc. of SPIE. 1995. Vol. 2419. P. 201—209.
12. Huang A.-M., Nguyen T. Correlation-Based Motion Vector Processing with Adaptive Interpolation Scheme for Motion-Compensated Frame Interpolation // IEEE Transact. on Image Proc. 2009. Vol. 18, N 4. P. 740—752.
13. Braspenning R., De Haan G. True Motion Estimation using Feature Correspondences // Proc. SPIE. Visual Communications and Image Proc. 2004. Vol. 5308. P. 396—407.
14. Сэломон Д. Сжатие данных, изображений и звука. М.: Техносфера, 2004. 368 с.
15. Yu L., Heindlmaier M. Optical Flow, Bilateral Filtering, Confidence Information and Motion Compensated Interpolation — a unifying approach. Technical report. Technische Universitat Munchen Institute for Data Processing. 2008. 99 p.
16. Orchard M. T., Sullivan G. J. Overlapped Block Motion Compensation: An Estimation-Theoretic Approach // IEEE Transact. on Image Processing. 1994. Vol. 3, N 5. P. 693—699.
17. Gilmutdinov M., Veselov A., Filippov B. Analysis of Side Information Generation Impact on Distributed Video Coding Performance // IEEE Proc. of XIII Intern. Symp. on Problems of Redundancy in Information and Control System. 2012. P. 26—31.
18. Набор тестовых видеопоследовательностей [Электронный ресурс]: <<http://media.xiph.org/video/derf/>>.

Сведения об авторах

- Антон Игоревич Веселов** — аспирант; Санкт-Петербургский государственный университет аэрокосмического приборостроения, кафедра инфокоммуникационных систем; E-mail: anton.veselov@gmail.com
- Марат Равилевич Гильмутдинов** — канд. техн. наук, доцент; Санкт-Петербургский государственный университет аэрокосмического приборостроения, кафедра инфокоммуникационных систем; E-mail: mgilmutdinov@gmail.com
- Борис Сергеевич Филиппов** — аспирант; Санкт-Петербургский государственный университет аэрокосмического приборостроения, кафедра инфокоммуникационных систем; E-mail: FilippovBoris@gmail.com

Рекомендована кафедрой
№ 51 безопасности информационных систем

Поступила в редакцию
01.02.13 г.