

Проблемы и перспективы развития систем кодирования динамических изображений

Виктор Дворкович, Михаил Чобану

Проводимые во всех технически развитых странах разработки алгоритмов и аппаратуры сокращения объема, рационального пакетирования и передачи по каналам связи с различной пропускной способностью видео-, аудио- и сопутствующей информации являются основой эффективного использования телекоммуникационных систем и радиочастотного спектра, сохранения действующих частотных планов, высвобождения значительной части частотного пространства для передачи потребителям дополнительных видов услуг – видеотелефонии, мобильной и стационарной видеоконференцсвязи, многопрограммного интерактивного телевидения, ТВЧ, трехмерного ТВ (3D-TV), телевидения ультравысокой четкости (ТУВЧ), многопрограммного звукового вещания, а также систем ТВ

с большим (в будущем – с произвольным) числом точек наблюдения [1-4].

Внедрение видеоинформационных систем в России связано с решением ряда проблем:

- ◆ эффективного использования мирового опыта по созданию и применению систем цифровой обработки и передачи информации и разработки соответствующих стандартов России;
- ◆ разработки и внедрения собственных мультимедийных систем и соответствующей аппаратуры;
- ◆ подготовки специалистов для обеспечения разработок, производства и эффективной эксплуатации таких систем;
- ◆ создания средств метрологического обеспечения цифровых видеоинформационных систем;
- ◆ интегрирования в общемировую систему телекоммуникаций с учетом тенденций перехода к системам

многопрограммного телевидения и телевидения высокой четкости в глобальном масштабе, передачи информации по наземным, спутниковым, кабельным и интернет-сетям.

Особое значение приобретает решение этих проблем в связи с разработкой «Концепции развития телерадиовещания в Российской Федерации на 2008–2015 гг.», одобренной распоряжением Правительства РФ №1700р от 29.11.2007.

В табл. 1 приведены параметры исходных и сжатых видеосигналов для приложений с одной точкой наблюдения, а в табл. 2 – для приложений со многими точками наблюдения. Как видно из этих таблиц, размер кадров изображений, их частота и формат существенно изменяются в зависимости от вида приложений. Величина компрессии

Таблица 1. Видеоинформационные системы с одной точкой наблюдения [1-4]

Приложения	Формат и размер кадра	Модель цветности и разрядность бит/пиксел	Частота кадров, Гц и тип развертки	Скорость несжатого потока, Мбит/с	Стандарты компрессии	Скорость сжатого потока, Мбит/с
Видеотелефония	subQCIF, QSIF, QCIF; 4:3	4:2:0; 12	1...5; прогрессивная	0,141...1,45	H.261, H.263, H.264, MPEG-4	0,009...0,128
Видеоконференцсвязь	QSIF, QCIF, SIF, CIF; 4:3	4:2:0; 12	5...30; прогрессивная	1,10...34,8	H.261, H.263, H.264, MPEG-4	0,032...1,50
Домашнее видео, видео по запросу	SIF, CIF; 4:3	4:2:0; 12	30/SIF, 25/CIF; прогрессивная	26,4/SIF; 29,0/CIF	MPEG-1/2/4	0,7...1,5
Потоковое видео в Интернете и беспроводных сетях 3G	QCIF, CIF; 4:3/16:9	4:2:0; 12	25, 30; чересстрочная	0,141...34,8	Фирменные, H.263, PEG-4, H.264/AVC, VC-1	0,02...0,6
Стандартное ТВ	640x480/NTSC	4:2:0; 12	30 – NTSC, 25 – PAL/SECAM; чересстрочная	105...141	MPEG-2, MPEG-4	2...20
	720x576 PAL/SECAM	4:2:2; 16		119...158		
ТВ повышенной четкости	1280x720; 16:9	4:2:0; 12 4:2:2; 16	25/30/50/60; прогрессивная	264...844	MPEG-2, MPEG-4	4...40
DVD, HD DVD, Blu-Ray	720x576; 4:3	4:2:0; 12	25/30, чересстрочная; 25, прогрессивная	10...949	MPEG-2, H.264/AVC, VC-1	1,858...9,8
	1920x1080; 16:9					10...20
ТВЧ 1080	1920x1080; 16:9	4:2:2; 16	25/30, чересстрочная	791...949	MPEG-2, MPEG-4, H.264/AVC	10...60
			25/30/50/60, прогрессивная	791...1900		
Цифровое кино	H1 1920x1080; 16:9	4:2:2; 20	60; прогрессивная	2370	MPEG-4, H.264/AVC	60...80
	H2 3840x2160; 16:9			9490		100...150
	H3 5760x3240; 16:9	38400		150...600		
	H4 7680x4320; 16:9	68300				
ТУВЧ	7680x4320; 16:9	4:4:4; 36	60; прогрессивная	72000	HEVC – High Efficiency Video Coding	100

Таблица 2. Видеоинформационные системы со многими точками наблюдения [5-8]

Многомерные телевизионные системы	Формат и размер кадра	Модель цветности и разрядность бит/пиксел	Частота кадров, Гц, тип развертки	Скорость несжатого потока, Гбит/с	Стандарты компрессии	Скорость сжатого потока, Гбит/с
Трехмерное ТВ – Side×Side; 2×1080i	1920×1080; (2×960)×1800; 16:9	4:2:0; 12	50/60; чересстрочная	0,8...0,95	MPEG-2, MP&HL,	0,065
					MPEG-4, AVC HP	0,035
Трехмерное ТВ – Blu-Ray; 3D(кино); 2×1080p	1920×1080; 16:9	4:2:2; 32	24; прогрессивная	1,58	MPEG-4/AVC Simulcast	0,05
					JPEG2000 Simulcast	0,24
Трехмерное ТВ – HD-3D; 2×1080p	1920×1080; 16:9	4:2:0; 24	50/60; прогрессивная	3,2...3,6	MPEG-2 Multiview Profile	0,095
Трехмерное ТВ – 2D+Depth: 1×1080i + карта глубины	1920×1080; 16:9	4:2:0; 12	50/60; чересстрочная	Нет данных	MPEG-C Part 3 Auxiliary Video Streams, MPEG-4 Part 2 Object Shape & Disparity Coding	Нет данных
Трехмерное ТВ – MultiView 3D-TV, n ракурсов; n×1080p	1920×1080; 16:9	4:2:0; 12	50/60; прогрессивная	n×1,5 ~ 100...1000	MPEG-4/H.264 AVC Multi-View Coding	0,100...1
Трехмерное ТВ – Free Viewpoint TV (FTV) – произвольное число ракурсов; n×1080p	1920×1080; 16:9	4:4:4; 24	50/60; прогрессивная	n×1,5 ~ 100...1000	MPEG-3DAV, MPEG-MVC, MPEF-FTV/3DV	Нет данных
Трехмерное ТВ – 3D-Integral Imaging, на основе интегральной фотографии; n2 ×4320p	7680×4320	4:4:4	60; прогрессивная	n2×1,5 ~ 15000	Нет данных	1...100
Голографические системы ТВ	4320×8192	Нет данных	60; прогрессивная	40x...1,5 ~ 10000...1000000	Нет данных	1...100

цифрового потока должна составлять от сотен (для систем с одной точкой наблюдения) до десятков тысяч раз (для многомерных систем).

В настоящее время наряду с улучшением существующих и созданием новых систем сжатия традиционных видеосигналов, снимаемых с одной точки наблюдения, бурно развивается как стереоскопическое (бинокулярное) телевидение, так и автостереоскопическое (с несколькими точками наблюдения, обычно 9...16 ракурсов). Также создано и развивается такое направление, как телевизионные системы с произвольной точкой наблюдения (ТСПТН или Free Viewpoint TV – FTV) [5-8].

На рис. 1 поясняются основные тенденции развития современных видеоинформационных систем. Наряду с ростом числа пикселей в системах ТВ с одной точкой наблюдения, выражающемся в переходе от систем ТВ к ТВЧ, и затем к ТУВЧ, наблюдается рост числа точек наблюдения – два ракурса для стереоскопических систем, до 16 ракурсов для систем многокурсового ТВ, и практически бесконечное число видов для систем ТСПТН (FTV).

Тенденции развития систем сжатия видеоинформации

Для характеристики кодирования видеоинформации можно использовать три фактора:

- ◆ оптимизированные методы сжатия, например, с использованием стандарта H.264/AVC, обладают выдающимися характеристиками, но приводят к связанным с потерями артефактам на определенных типах изображений; при этом тратятся существенные усилия на улучшение кодирования с целью уменьшения их визуальной зависимости;
- ◆ вычислительная сложность кодирования теперь становится важной проблемой, в том числе и из-за роста мобильных маломощных приложений, что в прошлом считалось не столь существенным фактором; это привело к необходимости учета ограничений на сложность алгоритма обработки;
- ◆ активно рассматриваются новые типы контента (трехмерное телевидение, многокурсовое видео и т.д.), при этом технологии сжатия развиваются так, что особое внимание уделяется общей эффективности кодирования, но зачастую оказывается неясным, как пользователи смогут получить доступ к этим

наборам данных, и насколько хорошо сжатые форматы подходят к типичным сценариям их использования.

Еще в 1929 году Рей Дэвис Келл предложил способ сжатия видео [9] путем передачи «только различия между последовательными изображениями объекта...». Этот способ все еще является краеугольным камнем многих стандартов сжатия динамических изображений.

Помимо устранения пространственной избыточности, эффективное временное предсказание движения деталей изображения и его компенсация всегда было ключевым фактором, определяющим общий коэффициент сжатия при кодировании видеоинформации [10].

Практически реализованные системы кодирования используют блочную оценку движения и его компенсацию, сопровождаемую блочным преобразованием изображения, квантованием и энтропийным кодированием. Даже после замены дискретного косинусного преобразования (ДКП) на его целочисленный вариант (в стандарте H.264/AVC) блочные технологии кодирования видеоинформации остаются подобными первоначально предложенным. Хотя очевидно, что «истинное» движение в кадре не является ни поступательным,

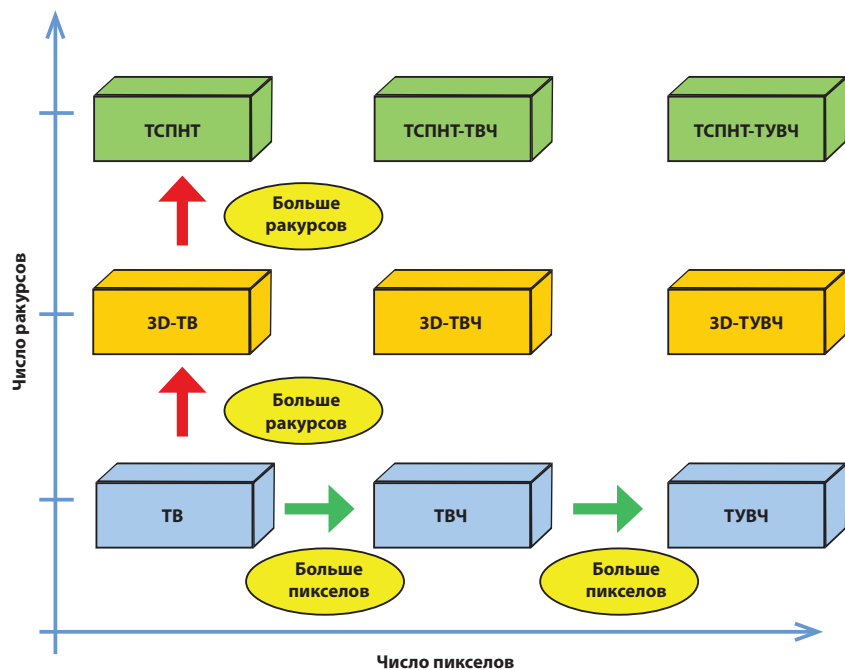


Рис. 1. Основные тенденции развития современных видеoinформационных систем

ни блочно-постоянным, блочная природа этих алгоритмов весьма полезна в вычислительном отношении. Последние разработки существенно не изменили способы оценки движения. Расширение коснулось изменения размеров блока, повышения пиксельной точности и новых режимов предсказания при внутрикадровом кодировании.

Значительные успехи были достигнуты в создании метрик качества, которые лучше оценивают субъективно (перцепционно) воспринимаемое качество. Однако в большинстве случаев эти новые метрики были разработаны при кодировании статических изображений [11]. Сопоставимые результаты для динамических изображений не так развиты. Развитие перцепционно ориентированных инструментов, которые включали бы временные качественные критерии, является ключевой задачей по улучшению рабочих характеристик систем кодирования видео. Заметим, что эти виды инструментов хорошо развиты для кодирования аудио.

По сравнению с ранее предложенными стандартами, H.264/AVC ввел многочисленные дополнительные режимы работы, включая большее число возможных векторов движения (например, векторы движения с 1/4-пиксельным разрешением), большее число размеров блока, несколько режимов внутрикадрового (intra) предсказания, и т.д.

Теоретически кодер будет более эффективным при применении соответствующих методов оптимизации,

реализация которых приводит к использованию множества способов выбора режимов его работы в зависимости от структуры кодируемых динамических изображений.

В связи с резким увеличением числа возможных комбинаций режимов и его влиянием на вычислительную сложность таких кодеков, необходимо проанализировать различные возможности построения пространства решений, исследовать алгоритмы и синтаксис кодирования, которые понизят общее количество допустимых режимов, ограничат выбор режимов блоков предсказания и компенсации движения деталей в нескольких кадрах динамических изображений [12].

Существуют приложения, в которых зрители имеют неполный доступ к части набора данных, в частности, телевизионные системы с произвольной точкой наблюдения (ТСПНТ – FTV). Такие системы требуют разработки кодеров, позволяющих декодерам обладать некоторой гибкостью. Предположим, сигналы от многих камер совместно кодировались с использованием инструментов, развитых в контексте кодирования множества ракурсов (MVC – multiview coding). Тогда пользователь может изменять угол наблюдения экрана, на самом деле выбирая только один из ракурсов в потоке бит. Новые наборы данных должны обеспечивать более полный доступ к видеoinформации, разрешая пользователям выбор точки наблюдения и управления произ-

вольным образом. При этом декодеру нужно позволить применять разные пути декодирования, каждый с различным порядком просмотра данных, таким образом соответствуя различным определениям того, что является будущим, а что – прошлым, с точки зрения цели декодирования.

Литература

1. **Дворкович В., Дворкович А.** Цифровые видеoinформационные системы (теория и практика) Части I, II. – М. НИИР-КОМ, 2010, 208 с.
2. **Дворкович В., Чобану М.** Телевидение послезавтрашнего дня // Радиочастотный спектр, 2010, №4, с. 26-31.
3. **Кривошеев М., Федунин В.** Видеoinформационные системы на основе телевидения высокой и сверхвысокой четкости // Broadcasting. Телевидение и радиовещание, №3, 2009.
4. **Дворкович В., Дворкович А.** Цифровые видеoinформационные системы в России. Современная электроника, 2008, №3, с. 8-13.
5. **Fujii T., Tanimoto M.** Free-Viewpoint TV (FTV) System. Lecture Notes in Computer Science, 2005, Volume 3331/2005, p. 497-504.
6. **Tanimoto M.** Free-Viewpoint Television. Image and Geometry Processing for 3-D Cinematography Geometry and Computing, 2010, Vol. 5, Part 1, p. 53-76.
7. **Kimata H., Kitahara M., Kamikura K., Yashima Y.** Free-viewpoint Video Communication Using Multi-view Video Coding. NTT Tech Rev/ Vol. 2, No.8, p. 21-26, 2004.
8. **Tanimoto M.** FTV (Free viewpoint TV) and Creation of Ray-Based Image Engineering. ECTI TRANSACTIONS ON ELECTRICAL ENG., ELECTRONICS, AND COMMUNICATIONS Vol. 7, No.2 August 2009, p. 1-12.
9. **Jacobs M., Probell J.** A Brief History of Video Coding. ARC International
10. **T. Koga, K. Iinuma, A. Hirano, Y. Iijima, and T. Ishiguro,** «Motion compensated interframe coding for video conferencing.» Proc. Nat. Telecommun. Conf, Vol. 5, No.3, p. 1–5, 1981.
11. **Z. Wang, A. Bovik,** «A universal image quality index», IEEE Signal Processing Letters, Vol. 9, No.3, p. 81–84, 2002.
12. **A. Ortega.** Video coding: Predictions are hard to make, especially about the future. Proc. Image Media Proc. Symp., 2007.

Продолжение следует