

А.Ю. АКСЕНОВ
**ИССЛЕДОВАНИЕ ПРИМЕНИМОСТИ СУЩЕСТВУЮЩИХ
МЕТОДОВ СЖАТИЯ К 3D-ВИДЕО ДАННЫМ**

Аксенов А.Ю. Исследование применимости существующих методов сжатия к 3D-видео данным.

Аннотация. В статье рассматривается проблема сжатия потока 3D-видео с использованием существующих видео кодеков. Проводятся сравнительные испытания существующих и предложенных методов.

Ключевые слова: 3D-видео данные, методы сжатия видеоданных, стереокодеки.

Aksenov A.J. The research of applicability of existing compression methods for 3D-video data.

Abstract. The article discusses the problem of compression of 3D-video data streams, using the existing well-known codecs. The comparison of the existing and proposed methods is giving.

Keywords: 3D-video, compression methods of video, stereo-codecs.

1. Введение. Основной поток информации об окружающей среде человек получает посредством зрения. Бинокулярное зрение — то есть видение двумя разнесенными в пространстве глазами, формирует ощущение глубины пространства. Из рис. 1 видно, что изображения окружающего мира на сетчатках обоих глаз не являются полностью одинаковыми.

Сложная биологическая система центральной (мозговой) части зрительного анализатора автоматически создает наиболее оптимальные условия для совместной работы обоих глаз, управляя глазодвигательными мышцами, диаметром каждого зрачка, конвергенцией глаза (сдвигом зрительных осей при зрении вблизи) и возвратом зрительных осей в параллельное состояние, когда человек перестает смотреть с близкого расстояния. Успех совместной работы обоих глаз выражается в образовании полноценного объемного зрительного образа, отражающего окружающий мир, в мозгу человека [1].

Изобретенный в 1838 году Чарльзом Уитстоном стереоскоп, заложил принцип формирования иллюзии трехмерного пространства с помощью пары двумерных изображений, полученных с разных ракурсов. С тех пор в мире накопилось большое количество стерео фотографий, как на бумаге, так и в цифровом виде. Тем не менее, до недавнего времени по причине технологических ограничений и высокой стоимости создания, дублирования и визуализации содержания в 3D технологиях не наблюдалось значимого прорыва. Появление цифро-

вых томографов еще в большей степени актуализирует проблему компрессии 3D-изображений [2].

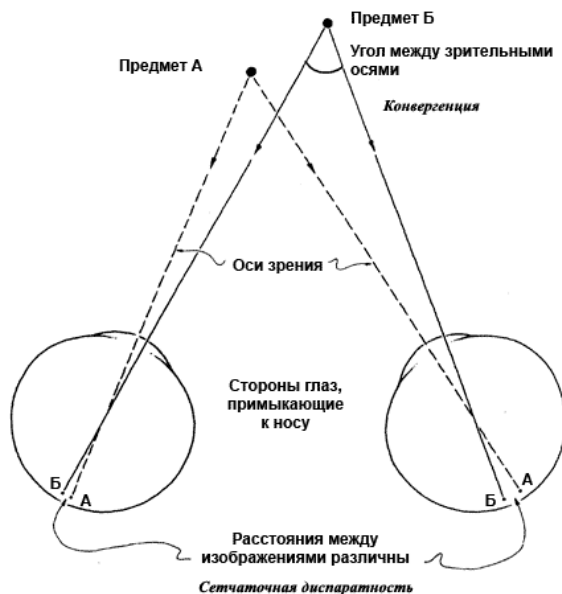


Рис. 1. Формирование отображения трехмерных объектов на сетчатках глаз.

С развитием цифровых технологий открылись широкие возможности массового создания, редактирования и распространения цифрового 3D-контента.

Однако сам принцип создания иллюзии объема по двум плоским ракурсам не изменился с XIX века. Это приводит к необходимости сохранения изображений с двух камер и, соответственно, удвоению объема хранимых данных. В связи с этим, становится актуальной проблема разработки форматов хранения и способов кодирования трехмерных сцен.

2. Сравнение методов сжатия 3D-контента. Известно, что современные видео кодеки хорошо справляются с устранением избыточности как внутри одного кадра, так и между следующими друг за другом изображениями одной и той же сцены. Так же известно, что смещение проекций трехмерных объектов, взятых с разных ракурсов, за-

висит от расстояния до объекта, и часть областей и элементов будут одинаковыми на изображениях, предназначенных для левого и правого глаза. Это создает предпосылки к проверке применимости современных двумерных видео кодеков для сжатия трехмерного контента.

Рассмотрим следующие варианты подготовки 3D контента для последующего сжатия:

1. Изображения левого и правого ракурсов размещаются в одном увеличенном кадре (рис. 2). Данный вариант является наиболее простым с точки зрения технической реализации, так как не требует сложного алгоритма преобразования. Кроме того, такой вид представления является одним из стандартных форматов для существующих устройств отображения (телевизоров и проекторов).

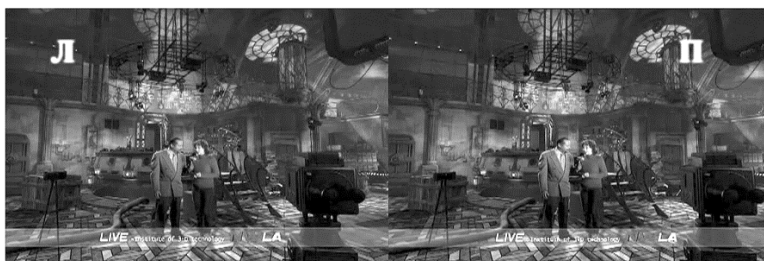


Рис. 2. Изображения левого и правого ракурсов размещаются в одном увеличенном кадре.

2. Формирование видео-потока, состоящего из чередующихся кадров левого и правого ракурсов (рис. 3). Этот вариант также прост с точки зрения технической реализации, но качественно меняет режим работы кодека, т.к. следующие друг за другом кадры существенно отличаются друг от друга.
3. Раздельное кодирование последовательностей кадров левого и правого ракурса (рис. 4). В этом случае на выходе кодека формируются два независимых потока для левого и правого глаза. Кодек работает в обычном режиме, а размер выходного потока в среднем оказывается в два раза больше чем для 2D варианта.



Рис. 3. Иллюстрация видео-потока, состоящего из чередующихся кадров левого и правого ракурсов.

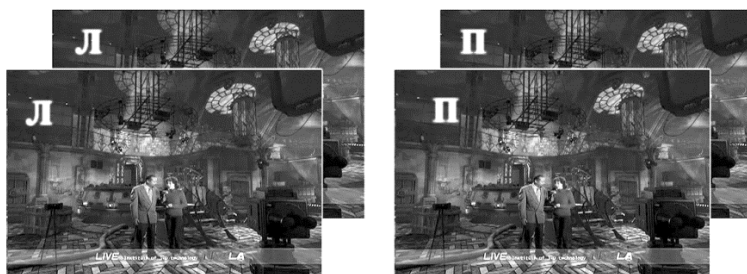


Рис. 4. Иллюстрация раздельного кодирования последовательностей кадров левого и правого ракурса.

4. Использование общих элементов в кадрах для правого и левого глаза. Специалисты Motion Picture Experts Group (MPEG) смогли использовать этот факт, чтобы снизить общий поток и размер файлов стереоскопических 3D-фильмов. Для формата Blu-ray 3D на основе кодека AVC (H.264) был разработан новый видеокодек MVC (Multi-View Codec), удерживающий увеличение потока (по сравнению со стандартным Blu-ray) на уровне 50%. Если пиковый поток для стандартных фильмов Blu-ray составляет 40 Мбит/с, то в случае Blu-ray 3D он увели-

чился до 60 Мбит/с. Blu-ray 3D MVC кодируется в виде основного потока видео (для одного глаза, то есть для воспроизведения в 2D) и дополнительного (зависимого) потока для другого глаза. Во втором потоке имеются ссылки на объекты в кадрах основного потока, то есть кодируются только различия [3, 4].

Для оценки эффективности использования каждого конкретного метода подготовки 3D видео потока для сжатия было проведено сравнение получающихся на выходе битовых объемов.

Результаты вариантов 1 и 3 легко предсказуемы и следуют из способа представления входных данных. Вариант 4 увеличивает объем 3D потока в среднем на 50% по сравнению с 2D вариантом. Наибольший интерес представляет второй вариант представления 3D видео-потока, позволяющий оценить применимость распространенных кодеков для сжатия нестандартных для них данных.

Для экспериментальной проверки из тестовой стерео видео пары формируется единая последовательность чередующихся кадров для левого и правого глаза, которая подается на вход кодеков h264 и Xvid.

Для удобства сравнения в одной из входных тестовых моно-последовательностей использовано значение битрейта (скорости битового потока) в 2 раза меньшее, чем в исходной стерео-последовательности.

Результаты экспериментов приведены в таблице.

Таблица. Результаты экспериментов

	Кодек H.264	Кодек XVID
Стерео (битрейт 4,5 Mbit/s)	69,9 Mb	72,8 Mb
Моно (битрейт 4,5 Mbit/s)	37,1 Mb	36,0 Mb
Моно (битрейт 2,25 Mbit/s)	18,7 Mb	18,2 Mb

Таблица иллюстрирует двукратное увеличение битрейта для потока чередующихся левого и правого ракурсов по сравнению с суммарным потоком, требующимся для представления двух независимых ракурсов. В соответствии с этим можно утверждать о невозможности применения видео-потока, состоящего из чередующихся кадров левого и правого ракурса (варианта 2) для представления стерео-видеопотока, что обуславливается внутренними механизмами представления последовательных кадров в видеокодеках H.264 и XVID, связанными с ис-

пользованием предсказателей движений и определением ключевых кадров.

При разработке стереокодексов следует учитывать особенности перемещений проекций объектов, которые выражаются не только в перемещении объектов в плоскости кадров, но и в изменении относительного смещения областей кадров, соответствующих перемещающимся в глубину объектам.

3. Заключение. Проведенные эксперименты показали ограниченную применимость существующих методов сжатия двумерных видеоданных, а также их модифицированных вариантов к стереоданным. Наиболее эффективным представляется разработка и использование специализированных ориентированных на 3D-контент кодеков, устраняющих, помимо внутри- и межкадровой избыточности, также избыточность, обусловленную схожестью кадров для левого и правого глаза.

Литература

1. *Лоточук Н.Н.* Принципы объемного видения // Электронный ресурс. — Доступ: <http://www.rudolf-steiner.ru/50000231/1529.html>
2. *Кулешов С.В.* Метод 3D-компрессии данных рентгеновской компьютерной томографии // Изв. вузов. Приборостроение. 2011. Т. 54, № 1, с.12–16.
3. Электронный ресурс. — Доступ: <http://total3d.ru/diy/92615/#.UTWZqtaePmc>
4. Extensions of H.264/AVC For Multiview Video Compression Emin Martinian, Alexander Behrens, Jun Xin, Anthony Vetro, Huifang Sun International Conference on Image Processing, October 2006. // Электронный ресурс. — Доступ: www.merl.com/papers/docs/TR2006-048.pdf

Аксенов Алексей Юрьевич — мл. научн. сотр. лаборатории автоматизации научных исследований СПИИРАН. Область научных интересов: обработка сигналов, обработка 3D-видео данных. Число научных публикаций — 10. a_aksenov@mail.iias.spb.su, sial.iias.spb.su; СПИИРАН, 14-я линия В.О., 39, Санкт-Петербург, 199178, РФ; р.т. +7(812)323-5139, факс +7(812)328-4450.

Aksenov Alexey Jurievich — Junior Reseacher, Laboratory of Research Automation SPIIRAS. Research interest: signal processing, 3D-video data streams processing. The number of publication — 10. a_aksenov@mail.iias.spb.su, sial.iias.spb.su; SPIIRAS, 14-th Line V.O., 39, St. Petersburg, 199178, Russia; office phone +7(812)323-5139, fax +7(812)328-4450.

Рекомендовано лабораторией автоматизации научных исследований СПИИРАН.
Статья поступила в редакцию 13.03.2013

РЕФЕРАТ

Аксенов А.Ю. Исследование применимости существующих методов сжатия к 3D-видео данным.

Основной поток информации об окружающей среде человек получает посредством бинокулярного зрения, которое формирует ощущение глубины пространства.

С развитием цифровых технологий появились возможности для массового распространения, обработки и создания 3D-контента.

Современные видео кодеки хорошо справляются с устранением избыточности как внутри одного кадра, так и между следующими друг за другом изображениями одной и той же сцены. Также известно, что смещение проекций трехмерных объектов взятых с разных ракурсов зависит от расстояния до объекта, и часть областей и элементов будут одинаковыми на изображениях, предназначенных для левого и правого глаза. Это создает предпосылки к проверке применимости современных двумерных видео кодеков для сжатия трехмерного контента.

В статье рассматриваются следующие варианты подготовки 3D контента для последующего сжатия: изображения левого и правого ракурсов размещаются в одном увеличенном кадре, формирование видеопотока, состоящего из чередующихся кадров левого и правого ракурсов, раздельное кодирование последовательностей кадров левого и правого ракурса.

Для экспериментальной проверки из тестовой стерео видео пары формируется единая последовательность чередующихся кадров для левого и правого глаза, которая подается на вход кодеков h264 и Xvid.

Для оценки эффективности использования каждого конкретного метода подготовки 3D видео потока для сжатия было проведено сравнение получающихся на выходе битовых объемов.

Проведенные эксперименты показали ограниченную применимость существующих методов сжатия двумерных видеоданных, а также их модифицированных вариантов к стереоданным.

Наиболее эффективным представляется разработка и использование специализированных ориентированных на 3D-контент кодеков, устраняющих, помимо внутри- и межкадровой избыточности, также избыточность, обусловленную схожестью кадров для левого и правого глаза.

SUMMARY

Aksenov A.J. **The research of applicability of existing compression methods for 3D-video data.**

The main source of information for humans is binocular vision which creates feel of depth of surrounding world.

With the development of digital technologies the wide possibilities for mass creation, processing and spreading of 3D-content has emerged.

The modern 2D video codecs are developed to decrease redundancy within the frame and between neighbor frames.

The frames for left and right eye in 3D-video sequences contain similarities which potentially can be reduced by 2D video codecs.

The paper considers variants of 3D video content preparation to be suitable for encoding by 2D codecs: the left and right images are packed in one bigger frame, left and right images are following one after another in one stream, left and right images are encoded into two independent streams.

For experiment testing the output of various frame packing methods was passed to h.264 and Xvid codecs for further processing. To estimate the efficiency of 3D-content preparation for encoding the resulting size of files were compared.

The experiments has shown the low efficiency of 2D video codecs for 3D content encoding. As a result the most progressive way to encode 3D video content is a development of special codecs which are to eliminate inter and intra frame redundancy but to cope with redundancy caused by similarity of images for left and right eyes.