

РАЗРАБОТКА МЕТОДОВ КОМПРЕССИИ АУДИО И ВИДЕО ИНФОРМАЦИИ С ИСПОЛЬЗОВАНИЕМ ОПТИМИЗИРОВАННЫХ ДЕЛЬТА-ПРЕОБРАЗОВАНИЙ ВТОРОГО ПОРЯДКА

Кравченко П.П., Хусаинов Н.Ш.

Таганрогский государственный радиотехнический университет
пер. Некрасовский, 44, ГСП-17А, Таганрог, Россия, 347928, каф. математ. обеспечения и применения ЭВМ
т. (86344)6.1903, (86344)6.1673 E-mail: kravch@tsure.ru, naile@mopevm.tsure.ru

Реферат: В работе рассматриваются новые подходы к цифровой компрессии аудио и видео данных на основе алгоритмов оптимизированных дельта-преобразований второго порядка, характеризующихся простотой реализации, достаточно высоким качеством преобразования, устойчивостью. Проблема обеспечения достаточного качества кодирования решается путем использования этих алгоритмов в сочетании с оригинальными методиками коррекции ошибок. Приводятся результаты экспериментальных исследований разработанных методов, обсуждаются перспективы их применения в современных системах связи и телекоммуникаций.

1. Современные тенденции в области цифровой компрессии аудиовизуальной информации

Вопросы компактного представления аудио и видео информации с целью ее эффективного хранения и передачи привлекают внимание специалистов на протяжении уже нескольких десятилетий, причем, несмотря на существующее многообразие методов и алгоритмов компрессии, программных и аппаратных аудио и видео кодеков интерес к ним не ослабевает. Это объясняется ориентацией тех или иных методов кодирования на конкретные приложения уже на этапе разработки. С учетом постоянно повышающихся темпов внедрения средств вычислительной техники во все сферы человеческой деятельности а также увеличением объемов подлежащей хранению, обработке и передаче информации (в первую очередь – такой "удобной" для восприятия, как аудиосигналы и видеоизображения), можно говорить, что разработка методов аудио и видео компрессии на сегодняшний день является одним из наиболее перспективных направлений развития информационных технологий.

Особенностью цифрового представления аудио и видео сигналов является наличие в них не только статистических, но и психофизиологических избыточностей, эффективное устранение которых позволяет добиться более высокой степени сжатия и в то же время не оказывает существенного влияния на их восприятие человеком.

Основными характеристиками, оказывающими влияние на использование того или иного метода кодирования для целей определенного приложения, являются эффективность алгоритма кодирования (соотношение коэффициента сжатия и степени соответствия восстановленного сигнала исходному) и трудоемкость его реализации. Значительная часть известных подходов в данной области для достижения максимальной эффективности кодирования базируется на переводе и последующем сжатии сигналов в частотной области (с использованием Фурье-подобных преобразований в сочетании, в частности, с энтропийными методами). В то же время известна высокая трудоемкость такого рода алгоритмов, что влечет за собой высокие требования к вычислительной мощности не только кодирующего, но и декодирующего оборудования, а в случае их аппаратной реализации – высокую стоимость специализированных плат аудио и видео компрессии.

Широко используемые ранее алгоритмы компрессии сигналов на основе поэлементного кодирования, характеризующиеся простой реализацией и высоким быстродействием (ИКМ, дифференциальная ИКМ, дельта-модуляция) вследствие их низкой эффективности сжатия зачастую не могут обеспечить эффективное хранение видеоизображений и аудиосигналов или их передачу по имеющимся каналам связи.

В данной работе рассматривается подход к цифровому кодированию аудио сигналов и видеoinформации на основе оригинальных алгоритмов оптимизированных дельта-преобразований второго порядка. Применение этих алгоритмов позволяет обеспечить увеличение отношения сигнал/шум при кодировании сигналов, характеризуется слабой чувствительностью аппроксимирующей функции к кратковременным возмущениям (помехам) независимо от их амплитуды, оптимизацией по быстродействию обработки скачка (переходный процесс) модулируемой функции, а также по точности в конце этого процесса, алгоритмической простотой реализации, одновременным сглаживанием и сжатием функции за счет замены полноразрядных отсчетов исходного сигнала одноразрядными знаками квантов вторых разностей демодулированной функции [1,2].

Использование высокопроизводительных алгоритмов дельта-модуляции в сочетании со специальными методиками, позволяющими обеспечить достаточное качество кодирования сигналов (сравнительные оценки см. ниже), позволяет говорить о перспективности практического применения методов, основанных на поэлементном кодировании, и об их эффективности, приближающейся к методам кодирования с Фурье-подобными преобразованиями.

2. Цифровая компрессия видеоизображений на основе алгоритмов оптимизированных дельта-преобразований второго порядка

Использование теории оптимизированных дельта-преобразования второго порядка позволило получить работоспособные алгоритмы с высокой скоростью изменения аппроксимирующей кривой, что обуславливается многоразрядными первыми разностями демодулированной функции. Однако современные требования к качественным характеристикам хранимых и передаваемых изображений привели к необходимости использования в сочетании с алгоритмами дельта-преобразований дополнительных методик, обеспечивающих заданное качество кодирования.

Суть разработанного метода покадровой компрессии видеоизображений состоит в построчном кодировании участков изображения алгоритмами оптимизированных дельта-преобразований второго порядка с последующей коррекцией ошибок дельта-преобразования на участках, где эта ошибка превышает определенное предельно допустимое значение. Для коррекции используются алгоритмы, основанные на принципах усеченного блочного кодирования, также характеризующиеся высокой простотой и приспособленные для сохранения высокочастотной составляющей спектра ошибки.

Адаптация кодера к характеру исходного изображения позволили сократить объем выходного потока за счет анализа ошибки кодирования участка после каждого "прохода" кодера и применение следующего уровня кодирования только в случае необходимости.

При проведении экспериментальных исследований алгоритма кодирования изображений были получены коэффициенты сжатия в среднем порядка 11-13 раз и отношения сигнал/шум порядка 31-32 дБ без предварительной обработки изображения. При этом визуальное качество закодированных изображений было сопоставимо с международным стандартом JPEG (при разнице в среднем около 2-3 дБ в пользу JPEG). Некоторый проигрыш в степени сжатия (в среднем, около 1.1 раза) компенсировался существенным повышением быстродействия в первую очередь алгоритма восстановления изображения (в 2.7-3.0 раза).

Следует отметить, что приведенные результаты экспериментальных исследований были получены на наборе тестовых изображений, включающих графические данные различных типов: фотографии, таблицы, рисунки, текстуры.

3. Цифровая компрессия аудиосигналов на основе алгоритмов оптимизированных дельта-преобразований второго порядка

Наряду с технологиями, ориентированными на достижение высокого качества аудио кодирования, например, в звукозаписи (стандарт MPEG, Layer III) и отличающимися крайне высокими требованиями к вычислительной мощности, достаточно широкое распространение при компрессии аудиосигналов получили алгоритмы на основе поэлементного кодирования (ИКМ, ДИКМ и т.п.). Это объясняется широким использованием подобных схем в бытовых системах связи с более низкими требованиями к качеству кодирования, например, в телефонии, где одним из определяющих факторов является стоимость аппаратуры, а значит и трудоемкость алгоритмов обработки сигналов.

В разработанном методе сжатие аудиопотока достигается при поблочной обработке аудио сигнала алгоритмами оптимизированных дельта-преобразований второго порядка за счет замены полноразрядного представления ИКМ-отсчетов в виде последовательности одноразрядных знаков квантов модуляции. Для восстановления одного отсчета сигнала требуется порядка 2-х операций целочисленного сложения.

Особенностью разработанного подхода является возможность использования кодера как в режиме с постоянной скоростью выходного потока (при гарантированной пропускной способности канала связи), так и в режиме работы с адаптацией скорости выходного потока к фактической пропускной способности канала связи. Реализация режима работы кодера с переменной скоростью в многоканальной системе речевого уплотнения позволяет обеспечить предельно высокое качество кодирования с учетом пропускной способности канала связи и количества задействованных абонентских линий. Другой областью применения такого режима работы является разработка систем обмена звуковыми сообщениями в IP-сетях с коммутацией пакетов, где скорость выходного потока кодера адаптируется к текущей пропускной способности канала связи, оцениваемой по количеству дошедших до адресата пакетов.

Следует также отметить универсальность разработанного подхода, т.е. возможность кодирования на основе единого алгоритмического аппарата аудиосигналов различных типов и с различными частотами дискретизации: телефонная речь, широкополосная речь, музыкальный аудиосигнал.

Проведенные экспериментальные исследования позволили получить показатели словной разборчивости речевых сигналов от 87% (при скорости выходного потока кодера 8 Кбит/с), до 100% (при скорости 32 Кбит/с и выше). Сравнительные субъективные оценки качества кодирования по 5-бальной шкале MOS стандартных подходов и разработанного метода для различных типов аудиосигналов приведены в таблице 1 [3].

Таблица 1

Тип аудиосигнала, область использования	Схема кодирования	Частота дискретизации, КГц	Выходной поток (моно), Кбит/с	Битов на отсчет	Оценка MOS
речь, телефония	CELP	8	8.0	1.0	3.9
речь, телефония	разработанный метод	8	8.0	1.0	2.3
речь, телефония	CELP	8	16.0	2.0	4.1
речь, телефония	разработанный метод	8	16.0	2.0	4.0
речь, телефония	G.721 ADPCM	8	32.0	4.0	4.1
речь, телефония	разработанный метод	8	32.0	4.0	4.4
речь, радиовещание	G.722	16	48.0	3.0	3.7
речь, радиовещание	G.722	16	56.0	3.5	4.0
речь, радиовещание	G.722	16	64.0	4.0	4.1
речь, радиовещание	разработанный метод	22	22.0	1.0	4.1
речь, радиовещание	разработанный метод	22	44.0	2.0	4.5
музыкальный сигнал	MPEG Layer I	44	224.0	5.1	5.0
музыкальный сигнал	MPEG Layer II	44	192.0	4.4	5.0
музыкальный сигнал	MPEG Layer III	44	64.0	1.5	4.9
музыкальный сигнал	разработанный метод	44	44.0	1.0	2.0
музыкальный сигнал	разработанный метод	44	88.0	2.0	3.6
музыкальный сигнал	разработанный метод	44	176.0	4.0	4.7

4. Заключение

Простота алгоритмов обработки позволяет эффективно использовать предложенный авторами подход при разработке высокопроизводительных систем цифровой компрессии аудио- видеосигналов в реальном масштабе времени. Низкая стоимость и достаточно высокая эффективность таких кодеков позволяет говорить о перспективности их применения как в компьютерных системах, так и в системах видеосвязи, охраны, проведения видеоконференций, многоканального уплотнения абонентских линий в каналах телефонной связи, при разработке специализированных вычислительных модулей.

Библиография

1. Кравченко П.П. Основы теории оптимизированных дельта-преобразований второго порядка. Цифровое управление, сжатие и параллельная обработка информации: Монография. Таганрог: Изд-во ТРТУ, 1997.
2. Кравченко П.П., Хусаинов Н.Ш. Цифровое кодирование аудиовизуальной информации на основе алгоритмов дельта-преобразований второго порядка: Тез.докл. Вторая международная конференция DSPA-99. – Москва, МЦНТИ, 1999г.
3. Holl P. Digital Audio Coding for Visual Communications //Proceedings of the IEEE. – 1995. – v.83, № 6



DEVELOPMENT OF AUDIO AND VIDEO COMPRESSION TECHNIQUES BASED ON OPTIMIZED SECOND ORDER DELTA-TRANSFORMATIONS

Kravchenko P.P., Khussainov N.Sh.

Taganrog State University of Radio Engineering
Nekrasovsky St., 44, GSP-17A, Taganrog, Russia, 347928
Tel.: (86344)6-19-03, (86344)6-16-73 E-mail: kravch@tsure.ru, naile@mopevm.tsure.ru

Abstract. New approaches to digital audio and video compression based on stable, high-efficient, simple to implement algorithms of optimized second order delta transformations are proposed. The problem of supporting enough coding quality is solved by using of that algorithms in conjunction with original error correction techniques. The results of tests are given. Prospects of applying in telecommunications are discussed.

Modern trends in audio and video compression

The problems of audio and video bit rate reduction for purposes of their effective storage and transmission have being attracted some last decades. Though there are a variety compression techniques and software or hardware codecs this problem has being actual. It is explained by development of every compression scheme for particular applications. Penetrating of computing systems to all fields of human activity and increasing of information content (especially suitable for human perception such as audio signal and video pictures) allow to consider audio and video compression as one of the most perspective branch of information technologies.

The main properties of compression scheme are their measure bitrate/quality and complexity (computational requirements). The considerable proportion of known coding schemes are based on frequency domain coding (usually with using Fourier-like transformations in conjunction with entropy coding) to achieve the minimum of bitrate/quality ratio. It is known that high complexity of such approach leads to high computational requirements for coding and decoding or to high cost of specific hardware.

The known algorithms for pel-by-pel coding (PCM, ADPCM, delta-modulation) used before and featured by low complexity now are unable to provide cost-efficient audio and video storage and transmission due to their low compression efficiency.

The approach to digital audio and video coding based on original algorithms of optimized second order delta-transformations is considered in this paper. Applying of that algorithms allows to provide signal/noise ratio increasing is characterized by low sensitivity of approximate function to short-time disturbances independently of their amplitude, optimization of step response by speed (transient), optimization of accuracy at the end of transient, low complexity, integrated smoothing and compressing by means of substitution of full range source samples by single-bit digital representation of second order difference signs [1,2].

Digital video pictures compression technique based on optimized second order delta-transformations

Applying of the theory of optimized second order delta-transformations allowed to get serviceable algorithms with high speed variation of approximate function due to multi-bit first order differences. But high modern requirements to image quality causes using of special error correction techniques in conjunction with delta-transformations algorithms.

The essence of developed technique for video frame compression is row-by-row coding of image segments with delta-transformations algorithms using and following error correction on the segments with non-tolerable error. The coding techniques based on block truncation coding (BTC) are used for correction.

Coder adaptation to local image activity reduces bitrate through analysis of error after each step and conditional applying of next level coding.

The main results of comparative experiments are following: compression ratio is about 11-13 under signal/noise ratio about 31-32 dB and subjective quality of decompressed image very closed to one of JPEG. Non-significant loss of compression ratio by our approach is overcome by appreciable gain of performance (2.0-3.0 times as large).

Digital audio compression technique based on optimized second order delta-transformations

In proposed method audio stream compression is achieved by substitution of multi-bit source samples by one or more single-bit delta-modulation quants. Only about two integer additions are required for a sample decoding. The advantage of developed approach is possibility to transmit bit stream over channels of limited guaranteed capacity such as dial-up telecommunications channels and to support variable-rate coding in packet-oriented networks and in multichannel system of speech compression. It is necessary to mention about possibility of different type source audio coding (telephone speech, wideband speech, wideband audio) based on unified algorithmic scheme.

The main results of experiments are following: word intelligibility is from 87% (at bitrate 8 Kbit/s) to 100% (at bitrate 32 Kbit/s and more). Results of subjective quality tests are given in Table 1 [3] for comparative evaluation.

Table1

	Compression technique	Sampling Rate, KHz	Bitrate (single channel), Kbit/s	Bits per Sample	MOS
Telephone speech	CELP	8	16.0	2.0	4.1
Telephone speech	Proposed scheme	8	16.0	2.0	4.0
Telephone speech	G.721 ADPCM	8	32.0	4.0	4.1
Telephone speech	Proposed scheme	8	32.0	4.0	4.4
Wideband speech	G.722	16	48.0	3.0	3.7
Wideband speech	G.722	16	56.0	3.5	4.0
Wideband speech	G.722	16	64.0	4.0	4.1
Wideband speech	Proposed scheme	22	22.0	1.0	4.1
Wideband speech	Proposed scheme	22	44.0	2.0	4.5
Wideband audio	MPEG Layer I	44	224.0	5.1	5.0
Wideband audio	MPEG Layer II	44	192.0	4.4	5.0
Wideband audio	MPEG Layer III	44	64.0	1.5	4.9
Wideband audio	Proposed scheme	44	44.0	1.0	2.0
Wideband audio	Proposed scheme	44	88.0	2.0	3.6
Wideband audio	Proposed scheme	44	176.0	4.0	4.7

4. Conclusion

Simplicity of processing allows effective using of proposed approach in developing high-performance systems of audio- and video- digital compression in real time. Low cost and enough effectiveness of such codecs let them to find wide employing as in computing systems as in videocommunications, videoconferencing, security, multichannel compression in telephony, in developing specific compression hardware etc.

References

1. P.Kravchenko Bases of optimized second-ordered delta-transformations theory. Digital control, compression and parallel information handling. Monograph. Taganrog.-1997.
2. P.Kravchenko, N.Khussainov Digital audio and video coding based on algorithms of second-ordered delta-transformations: Paper in Proceedings of 2-nd International Conference DSPA -99. – Moscow, ISCTI, 1999.
3. P.Holl Digital Audio Coding for Visual Communications //Proceedings of the IEEE.-1995.-v.83, N. 6.