

Раздел IV. Новые информационные технологии

П.П. Кравченко, К.В. Погорелов, А.Н. Шкурко, Н.Ш. Хусаннов

РАЗРАБОТКА МОДЕРНИЗИРОВАННЫХ АЛГОРИТМОВ ВИДЕОКОМПРЕССИИ С ИСПОЛЬЗОВАНИЕМ ОПТИМИЗИРОВАННЫХ ДЕЛЬТА-ПРЕОБРАЗОВАНИЙ ВТОРОГО ПОРЯДКА ДЛЯ СИСТЕМЫ ВИДЕОКОНФЕРЕНЦСВЯЗИ "ДЕЛЬТА-КОНФЕРЕНЦИЯ"*

Переход от изолированных офисных и домашних персональных компьютеров к повсеместному использованию локальных и глобальных вычислительных сетей, повышение мобильности сотрудников современных компаний в числе прочих требований диктует необходимость организации оперативного взаимодействия между несколькими пользователями в реальном масштабе времени. Для решения этой задачи используются системы видеоконференцсвязи (ВКС), которые в настоящее время находят все более широкое применение в бизнесе, управлении, профессиональной подготовке, образовании. Их высокая эффективность по сравнению, например, с телефонной связью, факсом или электронной почтой объясняется возможностью аудиовизуального общения не только с одним, но и одновременно с несколькими собеседниками, а также возможностью предоставления дополнительных сервисов, таких как организация документ-конференции (виртуальной рабочей доски, чата, пересылки файлов).

К системам компьютерной ВКС принято относить системы, в которых обработку информации осуществляют персональные компьютеры (ПК). В набор оборудования для стандартных систем ВКС также обычно входят web-камера, микрофон и акустические системы (наушники). Сравнительно низкая стоимость "дополнительного" оборудования обуславливает высокую конкурентоспособность и перспективность подобных систем.

Необходимость одновременного и высокопроизводительного – в реальном масштабе времени – решения целого ряда задач (таких, как воспроизведение каждым ПК нескольких получаемых аудио и видеопотоков, кодирование и передача участникам сеанса связи своего аудиовидеопотока, организация сетевого взаимодействия с несколькими терминалами ВКС, согласованный контроль подключения и отключения участников сеанса, ограничение несанкционированного доступа к передаваемым данным, многосторонний фоновый обмен файлами и др.) обуславливает жесткие требования к обеспечению эффективности и быстродействия всех компонентов системы, в особенности к алгоритмам обработки (компрессии и деконпрессии) аудио и видеосигналов.

В ТРТУ в качестве опытного образца программной системы конференцсвязи с алгоритмами аудио и видеокомпрессии на основе оптимизированных дельта-преобразований второго порядка для локальных IP-сетей разработана программная система многоточечной ВКС "Дельта-конференция" для использования в локальных и корпоративных IP-сетях с коммутацией пакетов. На разработанную программную систему ВКС "Дельта-конференция" получено свидетельство об офици-

* Работа выполнена при поддержке РФФИ, грант № 05-07-08032

альной регистрации в Российском агентстве по патентам и товарным знакам (Роспатент). В основу метода кодирования видеоинформации в системе ВКС положены разработанные в ТРТУ алгоритмы оптимизированных дельта-преобразований второго порядка и алгоритмы разностного кодирования на основе шаблонных матриц. Отличительными особенностями разработанного метода, обеспечивающими целесообразность и перспективность его применения в программных системах передачи видеоданных с компрессией в реальном масштабе времени, являются:

- ◆ возможность обработки видеоинформации различной по своему содержанию и динамике изменения видеосцены на основе единого алгоритмического подхода;
- ◆ внутрикадровое сжатие и сглаживание сигнала на основе единого алгоритма дельта-преобразования второго порядка;
- ◆ межкадровое сжатие и отсечение "малоизменяемой" части сигнала на основе алгоритма кодирования с использованием шаблонных матриц;
- ◆ низкая трудоемкость алгоритмов компрессии и декомпрессии;
- ◆ возможность эффективного управления качеством кодирования за счет адаптивного изменения частоты поддискретизации исходных данных, требований к качеству кодирования видеоинформации и возможности гибкой подстройки временной поддискретизации (пропуск кадров при кодировании);
- ◆ представление сжатого битового потока кодера в виде, приспособленном для последующего эффективного криптографического преобразования.

Компрессия видеосигнала выполняется с использованием разработанных в ТРТУ алгоритмов оптимизированных дельта-преобразований второго порядка и разностного кодирования. Алгоритм дельта-преобразования второго порядка со сглаживанием, в котором первые разности представляются в виде многоразрядных, вторые – в виде одnorазрядных приращений. Кодирование видеопоследовательности может выполняться в двух режимах: с использованием только внутрикадрового кодирования, что позволяет уменьшить влияние потерь при передаче данных на качество воспроизводимого изображения; и с использованием внутри- и межкадрового кодирования, что позволяет существенно повысить степень сжатия видеосигнала и разгрузить канал связи при незначительном ухудшении качества "картинки".

Сущность разработанного метода внутрикадровой компрессии заключается в том, что кодируемое изображение после перевода в цветовое пространство YCbCr подвергается построчной обработке алгоритмом дельта-преобразования с компенсацией скачков. Ошибки перегрузки по крутизне дельта-преобразований устраняются алгоритмом компенсации скачкообразных изменений исходного сигнала, основанным на дифференциально-импульсном кодировании.

Для дополнительного повышения эффективности кодирования используется раздельная обработка составляющих изображения пониженной поддискретизации и остаточного изображения видеосигнала. Составляющая кадра с пониженной дискретизацией выделяется путем выделения в каждом микроблоке размером 2x2 пиксела левого верхнего пиксела. Полученное "прореженное" изображение обрабатывается алгоритмом оптимизированных дельта-преобразований второго порядка с дополнительной (при необходимости) компенсацией скачкообразных изменений сигнала с использованием ДИКМ. Оставшиеся пиксела изображения преобразуются в разности между исходным изображением и его закодированной "прореженной" составляющей, которые квантуются с использованием логарифмической шкалы квантования и затем сжимаются статистическим кодером.

Алгоритм ДИКМ для алгоритма компенсации скачков и для кодирования остаточной составляющей реализован по табличному принципу, с использованием отдельных таблиц для каждой из цветовых компонент изображения. Служебная информация, содержащая флаги необходимости коррекции, кодируется для блоков размером 2×2 .

Сущность метода межкадровой компрессии заключается в том, что при сжатии формируется разностный кадр, состоящий из непересекающихся блоков размером 4×4 пиксела, и к каждому блоку (при необходимости) применяется алгоритм кодирования с итеративным приближением (коррекцией) с использованием корректирующих матриц, состоящих из единичных коэффициентов и отражающих (путем чередования знаков коэффициентов) наиболее вероятные структуры блоков разностных кадров. Для кодирования блока выполняется поиск оптимальной матрицы из фиксированного множества (64 или 128 матриц) и масштабирующего множителя посредством вычисления свертки пикселей разностного блока и коэффициентов матрицы. В случае необходимости к оставшемуся после кодирования блоку ошибок вновь применяется данный алгоритм. Количество итераций (проходов) может зависеть от требуемой скорости выходного потока кодера. Теоретические оценки и результаты экспериментальных исследований подтвердили существенный выигрыш по быстродействию данного подхода по сравнению со стандартными методами видеоконпрессии.

Наличие в шаблонной матрице только значений "1" и "-1" позволяет существенно упростить алгоритм кодирования и декодирования, так как "сложная" операция умножения заменяется "простой" операцией изменения знака. Использование блоков размером 4×4 пиксела обосновано такими факторами, как особенности первичной фильтрации изображения в зрительном анализаторе, наличие таблиц типовых конфигураций и простота разработки новых таблиц, возможности эффективной реализации обработки в ММХ-регистрах (строка блока из 4-х элементов обрабатывается за одну операцию).

Для повышения эффективности кодирования применяется упрощенная процедура поиска оптимальных корректирующих матриц, путем выбора номера корректирующей матрицы исходя из пространственного размещения знаков ошибки внутри корректируемого блока. Это позволяет отказаться от процедуры поиска оптимальной матрицы по множеству корректирующих матриц и получать ее номер, например, табличным способом, что, естественно, приводит к существенному повышению быстродействия алгоритма межкадрового кодирования при незначительном ухудшении качества кодированного сигнала.

Обработка разностного кадра выполняется поблочно (размер блока 4×4 пиксела).

На первом этапе оценивается среднее изменение интенсивности сигнала между кадрами в пределах обрабатываемого блока. Если разброс значений текущего блока относительно их среднего арифметического $P1$ не превышает предельно допустимой границы, то блок не подвергается дальнейшему кодированию (фрагмент изображения не изменился) и в выходной поток записывается только среднее значение (с квантованием и последующим статистическим кодированием). В противном случае разности между средним значением и исходными значениями каждого пиксела блока S_{ij} подвергаются дополнительному кодированию с целью сокращения ошибки.

Алгоритм дополнительного кодирования разработан с учетом основных положений преобразования Уолша-Адамара, однако вместо использования полного набора базисных ортогональных матриц, состоящих из значений "1" и "-1", пред-

лагается использовать эвристически сформированный на основе экспериментальных исследований набор матриц (в общем случае необязательно ортогональных), отражающих наиболее часто встречающиеся структуры разностных блоков.

Наиболее близким по принципам реализации к предлагаемому методу является подход, в котором предлагается представление изображения в виде полных бинарных блоков размером 4x4 с последующим включением этих блоков в передаваемый поток (один бит на отсчет); в известных реализациях используется только один проход по матрице.

В целях повышения общей производительности программной реализации видеокodeка предлагается использование возможностей в области параллельной обработки данных на современных процессорах. На данный момент существует несколько различных производителей центральных процессоров общего назначения, из которых наиболее известными и распространенными являются фирмы Intel и AMD. Выпускаемые этими фирмами процессоры общего назначения поддерживают различные наборы инструкций, позволяющих вести параллельную обработку однотипных данных. Это так называемые SIMD-расширения основного набора команд процессора (SIMD – Single Instruction - Multiple Data – Одна Инструкция - Много Данных). Для процессоров фирмы Intel это наборы команд MMX, SSE и SSE2. А для процессоров фирмы AMD – 3DNow и MMX. Таким образом, общим для обоих производителей процессоров набором команд является MMX, позволяющий вести параллельные вычисления над несколькими операндами (до 8 операций за одну инструкцию).

На рис.1 показана структурная схема кодека видеопоследовательностей и взаимодействия его модулей. Возле каждого функционального блока указано приблизительное потребление данным модулем процессорного времени в процентах от общего времени работы модуля видеокodeка. Соотношения потреблений процессорного времени приведены для последовательности из 15 кадров (1 опорный и 14 промежуточных).

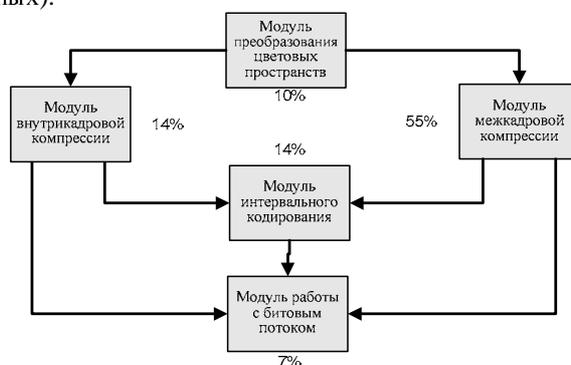


Рис.1. Структурная схема кодека видеопоследовательностей и взаимодействия его модулей

Использование в программной реализации видеокodeка набора распараллеливающих вычисления команд возможно во всех приведенных на рис. 1 функциональных модулях.

Рассмотрим теоретические возможности по распараллеливанию вычислений на различных этапах видеокodирования и их практические реализации с использованием набора инструкций MMX:

- ◆ параллельная реализация алгоритма преобразования цветových пространств. В модуле преобразования цветových пространств используется MMX-оптимизация основной процедуры преобразования цветов из одного пространства в другое. Для дополнительного повышения эффективности работы процедур преобразования используется табличное вычисление цветových компонент, при котором вместо 3-х операций умножения цветовой координаты на коэффициенты преобразования используется выбор результата умножения из заранее рассчитанных таблиц.
- ◆ параллельная реализация алгоритма работы с битовым потоком. В модуле работы с битовым потоком MMX-операции применяются для ускорения операций доступа для чтения и записи битового потока. Ускорение возможно за счет использования расширенных версий команд загрузки данных из памяти и сохранения данных в память. Также используются команды операций с битовым представлением чисел (операции сдвигов, логического "И" и "ИЛИ");
- ◆ параллельная реализация алгоритма работы модуля интервального кодирования. Модуль интервального кодирования использует MMX-операции целочисленных вычислений для чисел двойной разрядности (128 бит), что позволяет повысить на (1-2)% средний коэффициент статистического сжатия данных при незначительном ускорении процесса кодирования и декодирования информации;
- ◆ параллельная реализация алгоритма внутрикадровой компрессии. В модуле внутрикадровой компрессии повышение скорости кодирования возможно за счет применения команд групповой обработки данных на стадиях определения возможности применения локальной адаптивной поддискретизации, и при выполнении локальной пространственной поддискретизации изображения.
- ◆ параллельная реализация алгоритма межкадровой компрессии. В модуле межкадровой компрессии повышение скорости кодирования и декодирования возможно за счет применения команд групповой обработки данных на стадиях построения разностного кадра и восстановления кадра видеопоследовательности из разностного кадра. Структурная схема метода вычисления значений пикселей разностного кадра приведена на рис.2. Структурная схема метода вычисления значений пикселей декодированного кадра приведена на рисунке 3.

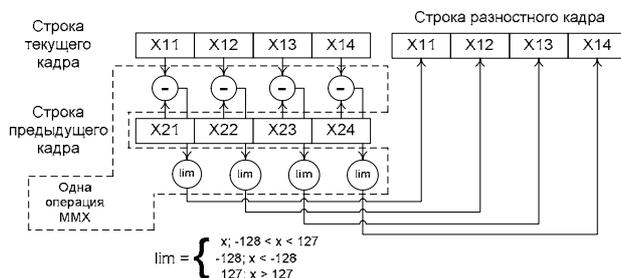


Рис.2. Структурная схема метода вычисления значений пикселей разностного кадра с использованием технологии MMX

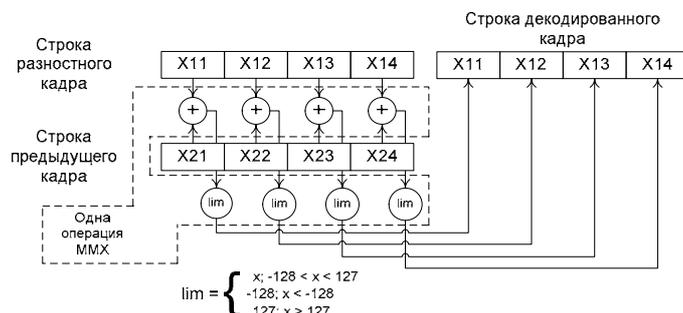


Рис.3. Структурная схема метода вычисления значений пикселей декодированного кадра с использованием технологии MMX

В соответствии с приведенными схемами оптимизации вычислений в программной модели видеокodeка разработана новая версия модуля сжатия для разрабатываемой системы ВКС. Результаты проведенных экспериментальных исследований влияния MMX-оптимизации на общую скорость работы видеокodeка приведены в таблице.

Таблица
Сравнение быстродействия разработанных видеокodeков без MMX-оптимизации и с MMX-оптимизацией с расчетом выигрыша в скорости обработки (ВСО) видеoinформации

Видео-последовательность	Разработанный кодек (с MMX-оптимизацией)			
	Скорость кодирования, кадров/сек (ВСО)	Средняя скорость выходного потока (коэф. сжатия)	Пиковая скорость декодирования с выводом на экран, кадров/сек. (ВСО)	Пиковая скорость декодирования без вывода на экран, кадров/сек. (ВСО)
160x120, 10 кадров/сек, 4,6 Мбит/с	135,0 (11,5%)	127,4 Кбит/с (36,1 раз)	505,6 (9,2%)	771,0 (12,9%)
320x176, 25 кадров/сек, 33,8 Мбит/с	34,2 (15,3%)	1 Мбит/с (33,8 раз)	159,0 (14,3%)	236,8 (17,8%)
320x240, 30 кадров/сек, 55,3 Мбит/с	31,4 (22,0%)	1,3 Мбит/с (43,7 раз)	121,3 (17,8%)	176,0 (21,4%)
560x352, 25 кадров/сек, 118,3 Мбит/с	13,5 (27,2%)	2,66 Мбит/с (44,4 раз)	50,5 (25,5%)	74,8 (32,2%)

Таким образом, производительность разработанного видеокodeка в среднем в 2,2 раза (при кодировании), и в среднем в 2,5 раза (при декодировании) превосходят программный кодек MPEG-1 при примерно одинаковых показателях эффективности кодирования (Pentium IV 1,4 ГГц, 256 Мб, Windows XP).

Полученные результаты позволяют говорить о перспективности разработанных методик кодирования и передачи данных конференцсвязи по сети и возможности их использования при разработке систем передачи и хранения аудиовизуальной информации общего и специального назначения (в том числе с защитой от несанкционированного доступа), конференцсвязи и др.