

УДК 681.007.05

А.В. Липанов, А.Ю. Михайлов

Харьковский национальный университет радиоэлектроники, Харьков

МОДЕЛЬ ПРЕДСТАВЛЕНИЯ МУЛЬТИМЕДИА ДАННЫХ В СИСТЕМАХ ПЕРЕДАЧИ ИНФОРМАЦИИ

Проблема передачи информации по сетям связи в реальном времени была успешно решена в 20-м веке. Несмотря на то, что 21 век уже переживает первое десятилетие, и информационные технологии развиваются все возрастающими темпами, проблема увеличения скорости передачи мультимедийных данных остается актуальной и сегодня. Интернет все чаще используется для передачи различных мультимедиа данных – видеоинформации, аудиоданных, изображений и электронных мультимедиа документов. Без сомнения, наличие спутниковых сетей и выделенных линий в больших городах сглаживают эту проблему. Но там, где нет таких мощных телекоммуникационных ресурсов, используя существующие способы передачи мультимедиа информации, часто необходимо уменьшать качество и информативность передаваемых данных. В статье рассматривается один из способов решения такой проблемы, а именно предлагается модель представления мультимедийных данных при их передаче в реальном времени.

Ключевые слова: мультимедиа системы, передача данных, модель данных, видеоданные, графика, сжатие данных в реальном времени.

Введение

Существует большое количество работ, в которых рассматриваются вопросы сжатия мультимедиа данных. Большинство этих работ характеризуется усилиями, предпринимаемыми по увеличению эффективности сжатия данных [1 – 4]. Такой подход позволяет достичь высокой эффективности сжатия данных, но далеко не всегда дает наилучшие результаты при передаче данных между компьютерами в локальной сети или сети Internet, особенно в условиях, когда необходимо одновременно передавать разнородные данные, такие как видео, звук, изображения экрана компьютера, текст, файлы. Именно такие задачи ежедневно решаются в современных системах удаленной поддержки пользователей, системах конференцсвязи, системах администрирования серверов и других. Таким образом, необходимо наличие эффективных алгоритмов, способных обеспечить специальную обработку данных с последующим сжатием в реальном времени.

Основной материал

В данной статье предлагается разработанная математическая модель представления мультимедиа данных, которая может быть эффективно применена в системах передачи в реальном времени. В ходе исследований в области передачи мультимедиа данных был обнаружен ряд проблем, связанных с неэффективностью описания данных с использованием классических математических моделей, существующих на сегодняшний день. В результате, для решения этих проблем был разработан ряд подходов, позволяющих существенно улучшить такие важные показатели системы передачи мультимедиа данных реального времени, как степень сжатия и время сжатия данных.

Под потоком мультимедиа данных подразумевается совокупность следующих типов данных:

- разнородный динамически меняющийся визуальный сигнал, содержащий данные с изображениями деловой графики, медицинской графики, искусственной графики, фотографии высокого качества, воспроизводимую видеозапись, изображения, полученные от специализированных устройств;
- визуальный сигнал, полученный с помощью видеокамеры;
- звуковой сигнал цифровой аудиозаписи музыкального произведения;
- звуковой сигнал, содержащий звук, полученный с микрофона;
- сигнал, содержащий текстовые данные;
- сигнал, содержащий любые двоичные данные.

Передача такого набора разнородной информации (потока мультимедиа) по каналу системы передачи данных с ограниченной пропускной способностью характеризуется временем задержки передачи данных:

$$\begin{cases} T(\bar{x}) = T_{\text{обр.перед.}}(\bar{x}) + T_{\text{перед.}}(\bar{x}') + T_{\text{обр.приемн.}}(\bar{x}'); \\ \|\bar{x}'\| < \|\bar{x}\|; \\ T_{\text{перед.}}(\bar{x}') < T_{\text{перед.}}(\bar{x}); \\ T_{\text{обр.приемн.}}(\bar{x}') > T_{\text{обр.приемн.}}(\bar{x}), \end{cases} \quad (1)$$

где \bar{x} – вектор данных, полученный от источника данных; \bar{x}' – вектор данных после дополнительной обработки передатчиком данных; $T_{\text{обр.перед.}}(\bar{x})$ – время обработки данных передатчиком данных; $T_{\text{перед.}}(\bar{x})$ – время передачи данных на приемник системой передачи данных; $\|\cdot\|$ – норма, представляющая размер вектора данных в байтах.

Важной особенностью суммарной задержки передачи данных по сети является то, что задержка передачи данных увеличивается при увеличении размерности передаваемого вектора. Для сокращения времени этой задержки применяются различные методы дополнительной обработки этого вектора, снижающие его размерность (уменьшающие объем передаваемых данных). Дополнительная обработка занимает некоторое время и, тем самым, увеличивает суммарную задержку. Таким образом, можно ввести оценку целесообразности уменьшения объема передаваемых данных посредством дополнительной обработки:

$$\begin{cases} \Delta T_{\text{перед.}} = T_{\text{перед.}}(\bar{x}) - T_{\text{перед.}}(\bar{x}'); \\ T_{\text{доп.обр.}} = T_{\text{обр.приемн.}}(\bar{x}') + T_{\text{обр.перед.}}(\bar{x}); \\ T_{\text{доп.обр.}} \leq \Delta T_{\text{перед.}} \end{cases} \quad (2)$$

Это выражение представляет собой ограничение на допустимое время, суммарно затрачиваемое на такие операции, как сжатие передаваемых данных с потерей и без потери качества с применением стандартных алгоритмов, а также выполнение дополнительных процедур, направленных на уменьшение объема передаваемых данных.

При соблюдении условий выражения (2) выполняемые операции по сжатию данных не оказывают влияния на время реакции системы, поскольку все затраты времени компенсируются за счет времени, сэкономленного непосредственно при передаче данных по сети. Система передачи данных, построенная по такому принципу, предоставляет наилучшее качество обслуживания конечных пользователей, передающих и принимающих мультимедиа данные по сети Internet в реальном времени.

На сегодняшний день существует большое количество алгоритмов сжатия данных как с потерей качества, так и без потери качества. Анализ этих алгоритмов показывает, что развитие направления сжатия данных, в целом, ориентируется на увеличение эффективности сжатия данных. Под эффективностью сжатия данных понимают коэффициент уменьшения объема данных после выполнения сжатия:

$$Q_{\text{сж.}} = \|\bar{x}'\| / \|\bar{x}\|. \quad (3)$$

Применение существующих алгоритмов сжатия и моделей представления позволяет существенно сократить объем данных. В то же время большинство этих алгоритмов и моделей, при непосредственном применении на практике, приводят к нарушению условий выражения (2). Это делает их малоприменимыми для создания системы передачи мультимедиа данных в реальном времени.

Часто задачу передачи мультимедиа данных пытаются решить посредством наращивания потенциала алгоритмов сжатия данных. Как правило, этот подход не является успешными по причине того, что в нем не учитываются условия выражения (2). Именно поэтому существует значительное количество систем сжатия мультимедиа данных [6], из которых только несколько систем способны работать в

реальном времени. В компьютерной графике принято описывать изображение как двухмерную совокупность точек, имеющих определенную яркость. Соответственно, поток видео данных представляется как совокупность двухмерных изображений [5], отображаемых в реальном времени:

$$S = \{B(i, j, t)\}, \quad (4)$$

где B – функция яркости видеоданных; i – первая координата изображения некоторого кадра; j – вторая координата изображения некоторого кадра; t – момент времени, которому соответствует некоторый кадр изображения.

При классическом подходе к описанию видеопотока существует необходимость обрабатывать и передавать большое количество данных, фактически ненужных конечному пользователю системы. То же самое касается и любого другого вида данных: звука и, в некотором роде, текстовой информации.

Продемонстрируем это на следующем примере. Пусть необходимо передать по сети видеосигнал электронной презентации. Такая презентация, как правило, содержит в видеопотоке большое количество изображений. Каждое из этих изображений представляет для конечного пользователя различный интерес. Так, например, фоновый рисунок презентации менее интересен просматривающему ее человеку, чем изображение в содержимом презентации, а изображение текстовых данных более интересно, чем вспомогательные рисунки. Схемы и графики также содержат большее число значащих деталей, чем, например, фотографии. Классическое описание видеопотока (4) как совокупности изображений экрана, отображаемых в реальном времени, не позволяет должным образом разграничить различную природу данных, содержащихся в видеопотоке. Это приводит к возникновению следующих проблем.

Поскольку все данные считаются однородными, для их обработки выбирается наиболее универсальный алгоритм сжатия, дающий приблизительно одинаковые результаты для всех вариантов данных, поступающих на вход системы сжатия. Такой алгоритм, будучи универсальным, соответственно, не обеспечивает наилучших результатов в каждом частном случае. Например, применение наиболее популярного алгоритма сжатия видео – MPEG-4 – не дает хорошего соотношения коэффициента сжатия (3) и качества изображения при обработке изображений текстовых данных, таких как титры в художественном фильме.

При универсальном сжатии видеоданных возникает также проблема подбора характеристик (параметров) алгоритма сжатия. Данные различной природы (происхождения) требуют применения различных параметров сжатия, в виду того, что данные, имеющие меньшую ценность для конечного пользователя, могут быть обработаны более интенсивно (с применением алгоритмов сжатия с потерей качества), чем более ценные для пользователя данные.

Классическое описание видеопотока имеет вычислительную сложность, описываемую выражением

ем (5). Данное выражение позволяет производить расчет количества операций, которые необходимо выполнить для сжатия t секунд видео потока с применением некоторого алгоритма A .

$$P(t', A) = t \cdot w \cdot h \cdot f_{\text{сек}} \cdot P_{\text{средн}}(A), \quad (5)$$

где A – алгоритм сжатия данных; t – длительность видео потока; w – ширина изображения; h – высота изображения; $f_{\text{сек}}$ – количество кадров видео потока в секунду; $P_{\text{средн}}(A)$ – среднее количество операций, выполняемых алгоритмом сжатия в процессе обработки одного байта информации.

Выполнение такого количества операций в реальном времени является невозможным на практике с применением современной техники.

Другой важной проблемой является параллельная обработка (сжатие) данных, описываемых различными математическими моделями. На практике, звуковой сигнал в системах мультимедиа данных всегда сопутствует видеосигналу и другим видам сигналов. Эти сигналы синхронизированы между собой, обрабатываются и передаются в реальном времени по единому каналу передачи данных, то есть фактически они являются единым цифровым сигналом. Несмотря на это, большинство современных подходов обработки данных рассматривают сигналы различной природы не как единый сигнал, а как совокупность нескольких независимых сигналов. Этот подход имеет такие недостатки:

- при чрезмерном потреблении ресурсов одним из сигналов нарушается передача остальных видов сигналов, поскольку алгоритмы обработки данных каждого отдельного сигнала не учитывают процессы, происходящие при обработке других сигналов;

- усложняется синхронизация передаваемых сигналов – некоторому фрагменту видеосигнала соответствует определенный звуковой сигнал и текстовый сигнал. Все три сигнала воспроизводятся синхронно, поскольку они являются связанными смыслом содержащейся в них информации. В случае независимой передачи этих трех сигналов возможна ситуация потери синхронизации, что приводит к искажению смысла информации, содержащейся в сигнале;

- в некоторых отдельных случаях различные сигналы могут содержать взаимосвязанные данные. Тогда рассмотрение их как единого сигнала позволяет добиться большего коэффициента сжатия данных (3).

Перечисленные выше проблемы решаются, если считать, что видеоданные, звук и любые другие данные являются единым сигналом (а не совокупности независимых сигналов), имеющим некоторую единую математическую модель. При этом модель должна быть достаточно универсальной, чтобы учитывать и общие свойства различных типов сигналов, и одновременно позволять учитывать индивидуальные особенности специфических типов сигналов. Например, сигнал видеоконференции, с одной стороны, должен обрабатываться как единая совокупность аудио и видео сигнала, а с другой стороны, должна

существовать возможность применять функционально различные виды сжатия для видеоданных и аудио данных. В нашей работе была построена математическая модель, обладающая такими свойствами, что и позволило получить высокие показатели эффективности системы обработки данных. Эта математическая модель имеет одну важную особенность. В отличие от большинства существующих моделей, описывающих данные с точки зрения их логического представления в ЭВМ, в нашей модели данные описываются, прежде всего, с точки зрения восприятия их пользователем – человеком. Такое описание в большинстве случаев не является лучшим с точки зрения простоты формализации, однако оно оправдывает себя на практике, так как важной особенностью мультимедиа сигнала является то, что этот сигнал интерпретируется непосредственно человеком. Исходный сигнал представляет собой некоторую совокупность визуальных, звуковых и прочих образов, воспринимаемую различными органами чувств. При передаче таких данных, сигнал не обязательно должен быть передан идеально точно, а лишь настолько точно, чтобы это не исказило его восприятия человеком. Данные, воспринимаемые человеком с помощью органов чувств, семантически представляют собой некоторую совокупность образов, которая не обязательно должна идеально точно формализоваться структурами данных ЭВМ. При передаче мультимедиа данных, в большинстве случаев, важной является только передача неискаженных с точки зрения человеческого восприятия образов, которые возникают у пользователя при получении этих данных. При этом любые искажения данных, которые не влияют на восприятие, являются приемлемыми, не оказывая негативного влияния на показатели качества системы передачи данных, что позволяет унифицировать модель, описывающую принципы представления мультимедиа данных и тем самым избежать недостатков классического подхода, описанных выше.

Пусть U – множество цифровых сигналов, представляющих собой последовательность байт, непосредственно передаваемых по сети, а O – множество оригинальных сигналов, которые должны быть переданы. Выражение (6) представляет собой модель классического подхода к передаче мультимедиа данных по сети.

$$k : O \rightarrow U. \quad (6)$$

Тут отображение k представляет собой некоторое преобразование оригинальных сигналов, выполняемое с целью обеспечения их передачи по сети, например, преобразование сжатия данных, преобразование устранения шумов, и т.д. Подход, описываемый выражением (6), был расширен нами с целью решения проблем, упомянутых ранее в данной работе.

Пусть A – множество алгоритмов преобразования оригинальных сигналов. Вместо преобразования множества оригиналов сигналов во множество передаваемых по сети сигналов, нами выполняется совокупность из двух преобразований: преобразова-

ния исходного сигнала в совокупность чувственных образов и преобразования чувственных образов в цифровой сигнал, передаваемый по сети:

$$\left\{ \begin{array}{l} O = O_1 \cup O_2 \cup \dots \cup O_j; \\ a_i \in A; \\ a_i : O_j \rightarrow I_i; \\ I = I_1 \cup I_2 \cup \dots \cup I_i; \\ m_e : I \rightarrow I'; \\ k_i : I_i' \rightarrow U, \end{array} \right. \quad (7)$$

где O_i – подмножества (классы эквивалентности), содержащие сигналы одинакового происхождения; O – множество оригиналов сигналов; A – множество алгоритмов извлечения чувственных образов из оригиналов сигналов; a_i – алгоритм извлечения чувственных образов определенного вида; I – множество чувственных образов; m_e – алгоритм преобразования множества чувственных образов; k_i – алгоритм окончательной обработки данных, выполняемой непосредственно перед передачей данных по сети.

Множество оригиналов сигналов может быть разбито на подмножества (классы) O_j , содержащие сигналы определенного происхождения (подмножество звуковых сигналов, подмножество изображений, подмножество текстовых сообщений и т.д.). Алгоритмы извлечения образов a_i отображают некоторые подмножества O_j в подмножества образов, полученных с применением данного алгоритма. Каждому алгоритму извлечения образов соответствует некоторый алгоритм, обеспечивающий передачу данных по сети – k_i . Задача алгоритма извлечения образов состоит в построении на основе множества оригиналов сигналов некоторых множеств I_i , элементы которых отобраны таким образом, чтобы соответствующий алгоритм упрощения передачи данных k_i гарантировал максимально возможный выигрыш при его применении.

Этот подход также позволяет описывать сигналы без учета способа их представления в памяти ЭВМ. Если $I_i \cap I_j = \emptyset$, то имеется возможность алгоритмы k_i и k_j подобрать таким образом, что эти алгоритмы будут использовать различные способы логического представления данных. Это дает возможность оптимальным способом описывать передаваемые данные и получать дополнительный выигрыш при их сжатии.

Следует заметить, что элементы множества I представляют собой отдельные визуальные, аудио, текстовые и другие образы, полученные путем преобразования оригинального сигнала. Мультимедиа сигнал в каждый момент времени эквивалентен некоторому небольшому подмножеству I . Само множество I фактически представляет собой совокуп-

ность всех чувственных образов, которые возникают при обработке данного мультимедиа сигнала. Существует возможность ввести набор операций $M = \{m_1, m_2, \dots, m_i\}$, преобразующих эти образы. Преобразование образов может выполняться с различными целями – улучшения их восприятия, фильтрации помех, устранения ненужных данных, и т.д. Например, может встретиться ситуация, что некоторые из образов являются ненужными, поскольку они не несут смысловой нагрузки, имеющей ценность для конечного пользователя. Такие образы могут быть удалены с применением некоторой специализированной операции m_i , что сократит объем передаваемых данных. Также может оказаться целесообразным объединение нескольких образов, либо разбиение образа на несколько образов с целью улучшения показателей оценки (3) – коэффициента сжатия данных. Набор операций m_i составляется с учетом особенностей алгоритмов k_i , применяемых для каждого вида чувственных образов. Правильный подбор этого множества операций позволяет существенно уменьшить объем передаваемых данных.

Продемонстрируем превосходство такого подхода на реальном примере. Пусть существует необходимость передачи электронной презентации продолжительностью 10 минут, состоящей из изображений фотографического качества, изображений векторной графики и изображений текстовых данных. Пусть данная презентация сопровождается музыкальным сигналом, представляющим собой 20 секунд аудио фрагмент, повторяемый циклически. В случае классического подхода к передаче такого сигнала сложность выполняемых операций описывается выражением (5). Введем a_1 – алгоритм извлечения фотографических изображений, a_2 – алгоритм извлечения изображений векторной графики, a_3 – алгоритм извлечения изображений текста, a_4 – алгоритм извлечения аудио образов. В итоге видео поток будет преобразован в совокупность отдельных образов, классифицированную по четырем категориям, для каждой из которой будет применен специализированный алгоритм сжатия, обеспечивающий наилучшее сжатие объектов данной категории. Например, в качестве алгоритма k_1 можно использовать JPEG 2000, в качестве алгоритма k_2 – GIF, в качестве алгоритма k_3 – специализированный алгоритм сжатия изображений текста, в качестве алгоритма k_4 – некоторый известный алгоритм сжатия звука. Такой подход обеспечивает возможность учета индивидуальных особенностей каждой категории образов при выполнении сжатия, что обеспечивает существенный выигрыш в объеме передаваемых по сети данных. Также существенная экономия объема передаваемых данных и вычислительных ресурсов происходит и за счет сокращения

количества повторно передаваемых образов. Пусть некоторый визуальный образ присутствует на каждом из 50 подряд идущих кадров видеосигнала. Этот образ в модели (7) будет представлен только одним элементом множества I . Вместо повторной передачи данных, описывающих образ, имеется возможность осуществить передачу этих данных только один раз. Это позволяет существенно сэкономить ресурсы, относительно случая, если бы использовалась классическая модель (4). Если при этом учитывать условия выражения (2), то на практике удастся решить задачу передачи мультимедиа данных по сети (ограниченной пропускной способностью в реальном времени) даже в случаях, если пропускная способность сети является критически низкой.

Выводы

Разработанная нами модель мультимедиа сигнала (2) и (7) позволяет создавать различные схемы обработки мультимедиа данных при их передаче по сети в реальном времени. Эти схемы по всем своим характеристикам существенно превосходят стандартный подход, когда мультимедиа сигнал представляется как совокупность независимых сигналов, а основная его составляющая – видеопоток – представляется как совокупность изображений, отображаемых в отдельные моменты времени. Основой разработанного подхода является представление мультимедиа потока как совокупности чувственных образов, каждому из которых соответствуют некоторые правила обработки данных, а также выполнение преобразований над данными образами с целью оптимизации процесса передачи данных. Заметим, что правила, приводимые в данной статье, одинаково хорошо выполнимы для всех видов мультимедиа данных: видео, звука, потоков изображений, потоков текстовых данных. В дальнейшем будет производиться детализация

модели, описывающей поток мультимедиа данных, изучение природы различных классов изображений, звуковых и других сигналов, будут разрабатываться алгоритмы выбора оптимальных механизмов обработки этих данных при передаче по сети в реальном времени. Для демонстрации преимуществ данного подхода будет разработано программное обеспечение передачи мультимедиа данных и будет проведен сравнительный анализ различных способов обработки и представления данных.

Список литературы

1. Cherriman P. Error-rate Based Power controlled Multimode H.263 -Assisted Video Telephony / P. Cherriman, L. Hanzo // *IEEE Transactions on Vehicular Technology*. – September 1999. – Vol. 48. – P. 37-41.
2. Bracamonte J. VLSI systems for image compression. A power-consumption / image resolution trade-off approach / J. Bracamonte, M. Ansorge, F. Pellandini, // *The International Society for Optical Engineering*. – October 1996. – Vol. 2952. – P. 1237-1241.
3. Clark N. Taylor. Adaptive Image Compression for Wireless Multimedia Communication / Clark N. Taylor, Sujit Dey, // *Electrical and Computer Engineering University of California*. – San Diego, La Jolla, California, USA, 2007. – P. 157-162.
4. Fang-Chu Chen. Trend in Multimedia Compression Technology // Fang-Chu Chen // *Wireless and Optical Communications Conference*. – Taichung, Taiwan, 2008. – P. 61-67.
5. Пуятин Е.П. Обработка изображений в робототехнике / Е.П. Пуятин, С.И. Аверин. – М.: Машиностроение, 1990. – 319 с.
6. List of multimedia streaming systems. – [Электронный ресурс]. – Режим доступа к материалу: http://en.wikipedia.org/wiki/List_of_streaming_media_systems

Поступила в редколлегию 17.10.2008

Рецензент: д-р техн. наук, проф. В.П. Пуятин, Харьковский национальный университет радиоэлектроники, Харьков.

МОДЕЛЬ УСВЛЕННЄ МУЛЬТІМЕДІА ДАНИХ В СИСТЕМАХ ПЕРЕДАЧІ ІНФОРМАЦІЇ

О.В. Ліпанов, А.Ю. Міхайлов

Проблема передачі інформації по мережах зв'язку в реальному часі була успішно вирішена в 20-му сторіччі. Незважаючи на те, що 21 сторіччя вже перетинає перше десятиріччя, та інформаційні технології розвиваються все більш зростаючими темпами, проблема підвищення швидкості передачі мультимедійних даних лишається актуальною й сьогодні. Інтернет все частіше використовується для передачі різноманітних мультимедіа даних – відеоінформації, аудіоданих, зображень та електронних мультимедіа документів. Без сумніву, наявність супутникових мереж та відокремлених ліній в великих містах спрощують цю проблему. Але там де немає таких потужних телекомунікаційних ресурсів, використовуючи існуючі засоби передачі мультимедіа інформації, часто необхідно зменшувати якість та інформативність даних які передаються. В статті розглядається один з засобів вирішення цієї проблеми, а саме запропоновано модель уявлення мультимедійних даних при їх передачі в реальному часі.

Ключові слова: мультимедіа системи, передача даних, модель даних, відеодані, графіка, стиснення даних в реальному часі.

MODEL FOR MULTIMEDIA DATA REPRESENTATION IN SYSTEMS FOR INFORMATION TRANSFERRING

A.V. Lipanov, A.Y. Mikhailov

Problem for data transferring in communication networks successfully resolved in 20th century. Not considering that finishing first 10 years of 21th century and information technologies develops with fast speed problem for increasing speed of multimedia data transferring stay actually today. Internet widely used for transferring different multimedia data like video, audio data, images and electronically documents. With no hesitations existing of satellite networks and dedicated lines in big cities make this problem easier. But in other places where no exists high speed telecommunication resources using of existing ways for multimedia data transferring, frequently need decrease quality and information's of transmitted data. In article was reviewed one of the ways for resolving of this problem which contain in proposed model for representation of multimedia data during their transferring in real time.

Keywords: multimedia system, data transferring, data model, video data, graphics, data compression in real time.