

УДК 621.39

А.В. Хаханова,
кандидат технических наук,
Н.А. Харченко,
В.Н. Кривонос

АНАЛИЗ ОБРАБОТКИ ПРЕДСКАЗЫВАЕМЫХ КАДРОВ В ТЕХНОЛОГИИ СЖАТИЯ ТРАНСФОРМИРОВАННЫХ ИЗОБРАЖЕНИЙ

Рассмотрены основные этапы формирования предсказываемых P-кадров, используемых при сжатии видеoinформации, приведены основные методы устранения избыточности, варианты кодирования трансформированных изображений и их недостатки.

Ключевые слова: кодирование с предсказанием, видеопоследовательность, P-кадры, статистическое сжатие, трансформированные изображения.

Розглянуто основні етапи формування передбачуваних P-кадрів, що використовуються при стиску відеоінформації, наведено основні методи вилучення надмірності, варіанти кодування трансформованих зображень та їх недоліки.

Ключові слова: кодування з передбаченням, відео послідовність, P-кадри, статистичне стиснення, трансформовані зображення.

Basic stages of the forming of the predicted P-shots, used for the compression of video information, are examined; basic methods of the redundancy elimination as well as the variants of the transformed images encoding and their failings are resulted.

Keywords: encoding with a prediction, videosequence, P-shots, statistical compression, transformed images.

Основной сложностью при работе с видео являются большие объемы передаваемой информации. Видео формата 720x576 пикселей 25 кадров в секунду в системе RGB требует потока данных примерно в 240 Мбит/с. Это делает необходимым использовать различные технологии сжатия. Однако современные системы сжатия не обеспечивают требуемой эффективности и не позволяют в полной мере подстроиться под изменяющиеся характеристики сетей передачи данных [1–4]. Поэтому актуальная тематика исследований заключается в проведении анализа и выработки направлений для дальнейшего совершенствования систем сжатия потокового видео.

Видеопоследовательность обрабатывается в современных технологиях сжатия путем разделения всех кадров на три типа: кадры первого типа (I-кадры, опорные) сжимаются с использованием информации, имеющейся только в этих кадрах. Кадры второго типа (P-кадры, предсказанные) сжимаются с использованием предшествующих I- или P-кадров с помощью предсказывающего кодирования и компенсации движения. Кадры третьего типа (B-кадры) сжимаются с использованием двунаправленного предсказания, т.е. с привлечением предшествующих и последующих I- и P-кадров. Максимальная степень сжатия

обеспечивается в кадрах В-типа. Однако такие алгоритмы являются наиболее ресурсоемкими, что в случае работы с потоковой передачей может привести к значительным временным задержкам в системах передачи данных. Поэтому возможны варианты использования только первых двух типов кадров. В работе [2] проанализирована обработка опорных I-кадров. Для получения более полного представления о возможностях технологий сжатия требуется проанализировать возможности обработки трансформированных кадров Р-типа. Таким образом, цель статьи заключается в анализе обработки предсказываемых кадров в технологии сжатия трансформированных изображений.

В Р-кадрах устраняется два вида избыточности: пространственная и временная (подобие между соседними кадрами). Сокращение пространственной избыточности достигается кодированием с преобразованием, а для временной – кодированием с предсказанием и блоковой компенсации движения. Относительно опорных Р-кадры сжимаются сильнее за счет использования факта о подобии соседних кадров в видеопоследовательности. Фактически Р-кадры содержат изменения относительно предыдущего I- или Р-кадра.

В зависимости от кодировщика, количество Р-кадров в видеопоследовательности может быть либо фиксированной частоты, либо определяться автоматически по сложности видеоряда. При воспроизведении видеоряда, для восстановления текущего Р-кадра, необходимо восстановить все предшествующие ему Р-кадры и ближайший предшествующий I-кадр. Результат кодирования каждого Р-кадра, также как и в опорных кадрах, подвергается обработке алгоритмом статистического сжатия и вставляется в поток отдельно от I-кадров и других Р-кадров. Алгоритм кодирования Р-кадров приведен на рис. 1.

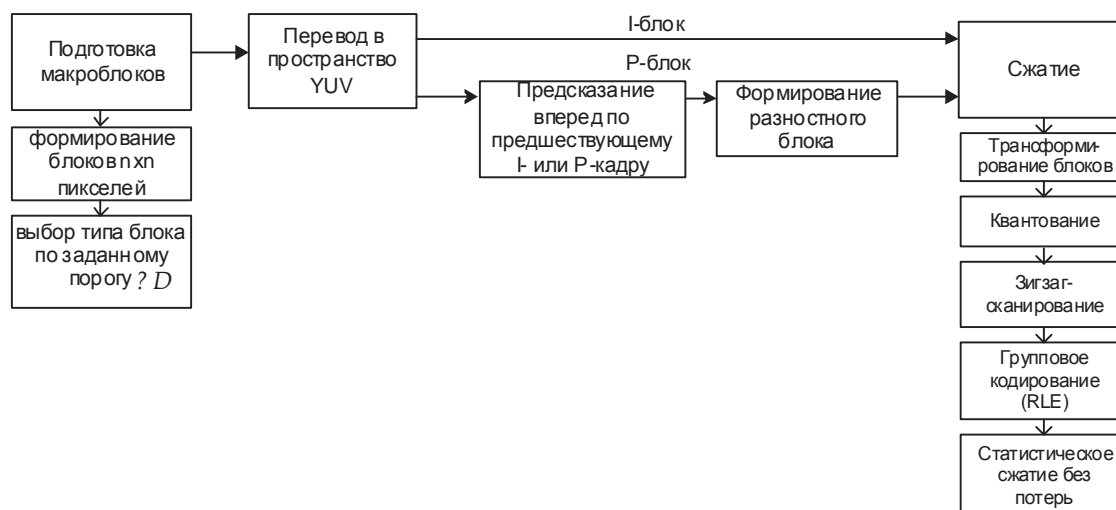


Рис. 1. Алгоритм кодирования Р-кадра

Весь конвейер преобразований можно представить с помощью следующих пунктов:

1. Подготовка макроблоков. Все изображение разбивается на блоки размером 8x8, 4 блока объединяются в 1 макроблок 16x16. Для каждого макроблока определяется режим сжатия. Производится сравнение блоков в текущем и предшествующем кадрах, и если разница не превышает заданного значения ΔD , то в

блоке используется кодирование с предсказанием (P-блок). Если в блоке появился новый объект и разница будет превышать ΔD , то происходит переход к алгоритмам, используемым для кодирования I-кадров. При выборе типа блока также необходимо учитывать три основных условия формирующих работу алгоритма сжатия: битовую скорость, вычислительную способность системы и требуемое качество рассматриваемого кадра. В соответствии с этими условиями выбирается как количество I-, P-кадров в видеопотоке, так и количество I-, P-блоков внутри P-кадра.

2. Перевод макроблока в цветовое пространство YUV. Преимущество пространства YUV по сравнению с RGB заключается в том, что цветоразностные компоненты U и V можно представлять с меньшим разрешением, чем яркостную Y.

3. Для P-блоков производится вычисление разности с соответствующим блоком в опорном кадре или кодирование с предсказанием. Простейшим вариантом кодирования P-блока является равномерное кодирование, при которой от значения уровня яркости одного отсчета отнимается уровень яркости последующего и кодируется их разность. Разностный блок формируется по формуле $\Delta x = x_{ij} - x'_{ij}$, где x_{ij} – значение пикселя текущего кадра с координатами i и j ; x'_{ij} – значение пикселя предшествующего кадра. На практике чаще используются методы кодирования разности отсчетов с предсказанием. На передающей стороне устанавливается предсказатель, которой по полученному в предыдущий момент отсчету $x(n)$ вырабатывает (предсказывает) последующий отсчет. При поступлении отсчета от передатчика предсказанное и истинное значение сравниваются (вычитаются) и передается ошибка предсказания $e(n)$. На выходе с учетом ошибок при передаче формируется $x'(n) \approx x(n)$.

Для предсказания формы сигнала, которая имеет достаточно прогнозируемые участки, используется кодирование с адаптивным предсказанием. Оно основывается на аппроксимации формы сигнала. Такая аппроксимация отображается следующей формулой:

$$x(T) = \sum_{t=1}^p a_t x(T-t), \quad (1)$$

где $x(T)$ – отсчет на выходе в следующий момент времени; a_t – коэффициент аппроксимации; p – порядок модели.

Важный принцип предсказания – это принцип “сохранения предыдущего значения”. В этом случае предполагается, что значение предыдущего отсчета сохранится и в последующий момент времени. Формула (1) тогда модифицируется следующей формулой:

$$x(T) = a_t x(T-1); \quad a_t = 1. \quad (2)$$

Для кодирования подается разность между текущим и предыдущим значениями:

$$\Delta x = x(T) - x(T-1). \quad (3)$$

Здесь величина $x(T)$ – аппроксимированное значение сигнала. Восстановление сигнала на приеме заключается в том, что к предыдущему значению сигнала прибавляется разность, $x(T) = x(T - 1) + \Delta x$. На рис. 2. показана одна из реализаций этого метода.

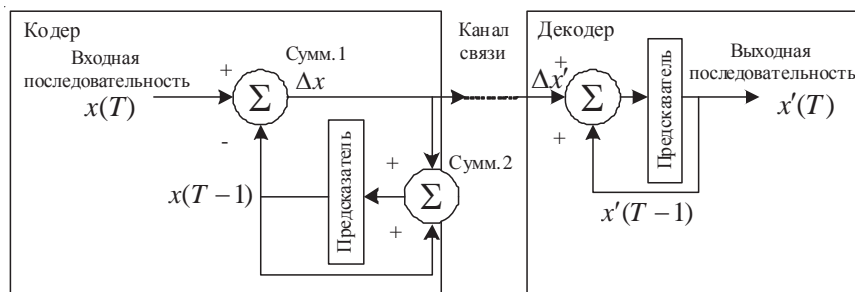


Рис. 2. Структурно-функциональная схема адаптивного кодирования с предсказанием

Разновидностью этого метода является аппроксимирование на основании нескольких отсчетов. Такой способ позволяет увеличить точность аппроксимации, но требует наличия памяти для накопления нескольких отсчетов. При предсказании, основанном на предположении “сохранение предыдущего значения”, коэффициенты не меняются, поэтому на приемном конце они те же самые – это позволяет не передавать их по каналу связи, а передать только разность аппроксимированного сигнала. Эффективность такого кодирования оценивается отношением η мощности сигнала, представленного кодами равномерного квантования, и мощностью сигналов кодирования разностного сигнала. Система становится неэффективной при большом значении разностного сигнала.

Потеря или искажение значения разностных отсчетов приводят к полному искажению восстанавливаемых значений, поскольку ошибка в предыдущем значении сигнала вызывает размножение ошибок в определении последующих. Поэтому все это требует передачи на приемный конец величины η выигрыша (для контроля качества принимаемого сигнала) и периодического обновления значений коэффициентов a_t (при устойчивом изменении характера сигнала). Это усложняет реализацию, которая показана на рис. 3 и рис. 4.

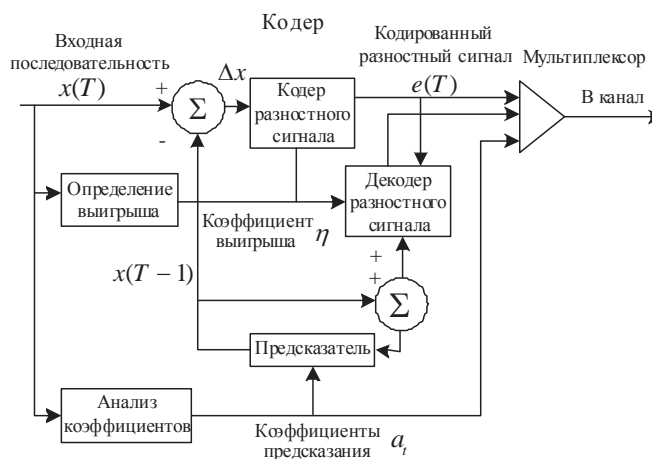


Рис. 3. Кодер с адаптивным предсказанием

На этих рисунках показаны кодер и декодер, которые способны передавать три блока информации: кодированный разностный сигнал $e(T)$, значение выигрыша η , коэффициенты аппроксимации a_t . Такая схема позволяет периодически проверять качество разностного кодирования и подстраивать коэффициенты аппроксимации. В этом смысле она адаптируема.

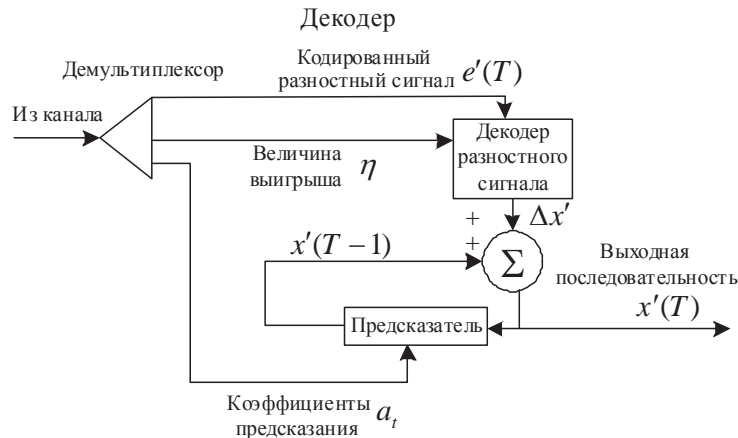


Рис. 4. Декодер с адаптивным предсказанием

Эффективность кодирования с адаптивным предсказанием зависит от сложности адаптивной логики и числа отсчетов для следующего предсказания. Существует оптимальная точка эффективности предсказания между большим накоплением статистики (осторожная тактика) и быстротой реакции на изменение (тактика быстрого реагирования).

4. Трансформирование блоков изображения. Трансформирование осуществляется с помощью дискретного косинусного преобразования (ДКП).

5. Равномерное скалярное квантование частотных коэффициентов. Процедура квантования заключается в сопоставлении номера кванта каждому символу трансформанты.

6. Зигзаг-сканирование или перевод трансформанты в вектор: $z_{ij} \rightarrow z_{[n]}$.

7. Групповое кодирование. Сжатие происходит за счет того, что в полученной кодовой группе встречаются цепочки одинаковых байт. Замена их на пары <счетчик повторений, значение> уменьшает избыточность данных. Однако такой тип кодирования будет эффективен только для черно-белого изображения, где вариантов цвета всего 2 и кодовая последовательность состоит из цепочек 0 и 1. Таким образом, чем больше цветовая гамма передаваемого изображения и чем больше на нем неоднородностей, тем меньше эффективность данного кода.

8. Статистическое сжатие без потерь (кодирование Хаффмана или арифметическое). Метод кодирования переменной длины сопоставляет потоку входных символов определенную последовательность кодовых слов (коды переменной длины, VLC – Variable Length Codes). Каждый символ отображается в кодовое слово, которое может иметь переменную длину, но каждое из них имеет целое количество бит. Эффект сжатия данных начинает проявляться после кодирования достаточно большого числа входных символов. Для достижения оптимального сжатия необходимо использовать разные таблицы для этих последовательностей, имеющих разные распределения вероятностей векторов.

Так как этот код характеризуется переменной длиной, то его применение требует буферного накопителя данных для выравнивания скорости передачи [4]. Таблицу вероятностей для длинной видеопоследовательности можно определить только после просмотра всей видеопоследовательности. Это может привести к недопустимой задержке процесса кодирования, передачи и воспроизведения видео на приемном конце. Также при формировании кода основываются на определенных предположениях о распределении вероятностей, характеризующем источник сообщений. В том случае, когда эти предположения не выполняются, в работе кодирующей системы может наступить резкое ухудшение, что в случае плохого согласования с источником кода может привести к увеличению средней затраты двоичных цифр.

Серьезным недостатком кодов, построенных на основе схемы Хаффмана, является их чувствительность к ошибкам. Искажение в одном бите последовательности этих кодов может привести к полной потери синхронизации при декодировании и к невозможности дальнейшего правильного декодирования последовательности.

Выводы

На эффективную всю систему сжатия значимое влияние оказывают предсказанные кадры (Р-кадры). Чем большее количество Р-кадров используется в видеопоследовательности, тем сильнее сжатие в целом, но тем хуже качество изображения и тем выше должен быть уровень контроля за ошибками. Ошибка в одном блоке распространится на всю группу последующих Р-кадров.

Кадры для технологий сжатия потокового видео являются зависимыми. Это обстоятельство накладывает условие неотделимости на кадры в пределах группы одного ключевого кадра.

Используемое в современных алгоритмах сжатия трансформированных изображений статистическое сжатие, основанное на кодировании Хаффмана, имеет ряд существенных недостатков, таких как: высокая чувствительность к ошибкам, необходимость в буферном накопителе, невозможность определить степень сжатия до начала кодирования, большие задержки при построении дерева для видеопоследовательностей, содержащих большое количество кадров.

Для увеличения степени сжатия видеоряда необходимо использовать несколько типов кадров с разной степенью сжатия, в которых будут устраняться различные виды избыточности.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. *Ватолин Д.* Методы сжатия данных. Устройство архиваторов, сжатие изображений и видео / Д. Ватолин, А. Ратушняк, М. Смирнов, В. Юкин. – М. : Диалог-Мифи, 2003. – 381 с.
2. *Ричардсон Я.* Видеокодирование. H.264 и MPEG-4 – стандарты нового поколения / Я. Ричардсон. – М. : Техносфера, 2005. – 368 с.
3. *Сэлмон Д.* Сжатие данных, изображений и звука / Д. Сэлмон. – М. : Техносфера, 2004. – 368 с.
4. *Баранник В.В.* Кодирование трансформированных изображений в инфокоммуникационных системах / В.В. Баранник, В.П. Поляков. – Х. : ХУПС, 2010. – 212 с.

Отримано 04.06.2012