



Международный журнал информационных технологий и энергоэффективности

Сайт журнала:

<http://www.openaccessscience.ru/index.php/ijcse/>



УДК 004.961

ОБЗОР СУЩЕСТВУЮЩИХ МЕТОДОВ СЖАТИЯ ВИДЕОДАНЫХ

¹Кирпичев Д.С., Шпедер Д.А, Трокоз Д.А..

ФГБОУ ВО «ПЕНЗЕНСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНОЛОГИЧЕСКИЙ УНИВЕРСИТЕТ», Пенза, Россия (440015, г.Пенза, проезд Байдукова, 1А/11), e-mail: ¹valerochkaleonich228@mail.ru

В статье описаны алгоритмы сжатия видеоданных с потерями. Выявлены ключевые проблемы, которые возникают при кодировании видеофайлов. Рассмотрены основные особенности и этапы сжатия видеопотока.

Ключевые слова: Сжатие, видеоданные, видеопоток, кодек, алгоритм, потери, кодирование, кадр, пиксел.

OVERVIEW OF EXISTING VIDEO DATA COMPRESSION METHODS

¹Kirpichev D.S., Shpeder D.A., Trokoz D.A.

PENZA STATE TECHNOLOGICAL UNIVERSITY, Penza, Russia (440015, Moscow, Vadkovsky Lane, 3A, p. 1), e-mail: ¹valerochkaleonich228@mail.ru

The paper describes lossy video compression algorithms. The main problems that arise when encoding video files have been identified. The main features and stages of video stream compression are considered.

Keywords: Video Data compression, video stream, codec, algorithm, lossy compression, coding, frame, pixel.

Введение

Необходимость качественной передачи информации с каждым днем возрастает. Современные каналы связи, по которым передаются видеоданные имеют ограниченную производительность, так как при передаче цифрового сигнала, он подвергается зашумлению. Кодирование информации обеспечивает компактное представление данных. Качество видеофайла зависит от следующих факторов: разрешение кадра, глубина цвета, частота кадров в секунду и алгоритм сжатия. Самое высокое качество видео из наиболее часто используемых пользователями на данный момент имеет разрешение 7680×4320 (33,2 мегапикселя), которое превосходит предыдущий стандарт телевидения сверхвысокой чёткости в четыре раза. С ростом качества передаваемой информации, растет и ее объем что повышает необходимость уменьшения размера данных. Желательным видом сжатия является сжатие без потерь, когда при раскодировании файла не ухудшается его качество. На практике такого качества достичь достаточно трудно, поэтому обычно применяется сжатие с потерями. При кодировании видеоданных с потерями полученный в результате сжатия файл будет иметь худшее качество относительно исходных данных. Сложность сжатия видеопотока состоит в необходимости работать с последовательностью кадров. При использовании методов сжатия видеопотока (ВП), исходный видеофайл рассматривается как перечень кадров, следующих друг за другом. [1] При сжатии видеофайлов необходимо учитывать, что порядок следования кадров должен

оставаться неизменным, так как если нарушить порядок кадров, то изменится и сам файл. Каждый кадр в процессе кодирования рассматривается как отдельное изображение, состоящее из пикселей.

Кодек – программный инструмент для кодирования видеоданных, под кодированием понимается процесс преобразования видеосигнала в наиболее удобный формат для хранения и передачи. Кодеки используются для сжатия и распаковки видео и аудио потока. Кодеки во время работы, в зависимости от их вида и цели кодирования, используют как метод сжатия с потерями, так и без потерь [2].

Задача сжатия видеоданных наиболее актуальна применительно к камерам видеонаблюдения, где в среднем видео, записанное одной камерой с разрешением записи 1920x1080 без применения кодеков сжатия занимает 20GB в минуту. В одной только Москве установлено 221 тысяч таких камер. Для своевременной и наиболее полной передачи информации необходимо использовать алгоритмы сжатия.

Сжатие с потерями

Данный вид сжатия можно разделить на следующие подгруппы: сжатие с естественными и неестественными потерями. В первом случае человеческий глаз не замечает дефектов изображения, полученных при сжатии видеопотока [3]. На восприятие объектов в окружающем мире оказывает влияние расстояние между объектом и угол его обзора, кроме того на изображение, которое получает человек, влияют различные помехи: дождь, туман и пр. Все эти факторы определяют уровень детализации изображений, воспринимаемых человеческим глазом. Орган зрения человека, устроен таким образом, что мы различаем острые грани и линии независимо от изменений в качестве изображения. [4].

При проигрывании видео, человек воспринимает изменения на экране монитора, как непрерывные, так как кадры меняются быстро. Для создания эффекта непрерывности достаточна смена десяти кадров за одну секунду. Все видеокодеки при сжатии разделяют видеофайл на кадры и работают с каждым изображением по отдельности. В компьютере изображения хранятся, как совокупности точек (пикселей). Цвет каждого пикселя. Для хранения и отображения изображения на электронных носителях используются цветовые модели. Наиболее популярной на данный момент цветовой моделью является модель RGB. Все цвета модели в данном случае формируются из трех цветов красного, зеленого и синего. По одному байту необходимо для кодирования красного зеленого и синего цвета, из этого следует, что для описания одного цвета из палитры необходимо три байта в памяти компьютера. При кодировании видеофайлов необходимо учитывать особенности хранения каждого кадра последовательности.

Далее приведем сравнительный анализ существующих методов сжатия видеофайлов.

Векторная квантизация

Для упрощения работы с изображением и сокращения объемов оперативной памяти, который необходимо затратить для работы с видеопотоком, необходимо разбить каждый кадр на блоки. [5] Сжатие выполняется посредством группировки блоков по общим признакам и замены группы схожих блоков одним большим; следующим шагом на основе данной замены формируется двоичная таблица (кодовая книга) общих блоков из наиболее коротких кодовых слов [6]. После этого из общих блоков собирается изображение, на основе полученной ранее

таблицы. Считывание элементов таблицы позволяет максимально сжать изображение с некоторыми потерями в качестве. Если уменьшить число классов и увеличить количество блоков, относительно которых определяется схожесть, то степень сжатия изображения будет максимальной. При уменьшении размеров таблицы значительно ухудшается качество исходного видео [7].

Возьмем три матрицы размерностью 4x4 (Рисунок 1). Данные матрицы представляют собой блоки, которые формируются кодеком при векторной квантизации [5].

(Блок 1)	(Блок 2)	(Блок 3)
128 128 128 128	128 127 128 128	128 127 126 128
128 128 128 128	128 128 128 128	128 128 128 128
128 128 128 128	128 128 127 128	127 128 128 128
128 128 128 128	128 128 128 128	128 128 128 128

Рисунок 1 – Сформированные матрицы

Для человеческого глаза эти блоки неотличимы, из этого следует, что 2-й и 3-й блоки можно заменить первым. В результате получаем кодовую книгу (Рисунок 2) [8].

128 128 128 128
128 128 128 128
128 128 128 128
128 128 128 128

Рисунок 2 – Кодовая книга

Дискретное косинусное преобразование (ДКП)

Одним из способов сжатия изображений является кодирование посредством преобразования. Данный способ применяется в таких форматах изображений как JPEG и в видеокодеках таких как MPEG-1, MPEG-2 и MPEG-4. Глаз человека может воспринять максимум 100 – 150 кадров в секунду, большая частота кадров имеет высокие показатели ДКП. Значения коэффициентов квантования прямо пропорционально зависят от частоты. Если рассматривать стандарт JPEG, то можно заметить, что чем выше частота коэффициентов ДКП, тем более высокие коэффициенты квантования необходимо применять. В данном случае используется математическая функция косинуса. Изображение делится на блоки 8 на 8 пикселей, каждый из этих блоков взвешивается на основе коэффициента – числа, которое определяет влияние блока на изображение в целом. Чем выше частота волны, тем более высокая частота сигнала. Косинусоидальные волны высокой частоты (цвет на изображении меняется резко) влияют на изображение в меньшей степени, чем волны низкой частоты и в процессе квантования принимают значения близкие к нулю и в процессе кодирования отсекаются. Далее происходит квантование, получившейся матрицы коэффициентов на

основе общепринятой матрицы квантования [9]. В данном случае выполняется квантование коэффициентов посредством кодирования кодов сигнала, в которых содержатся также и сигналы с нулевыми коэффициентами.

Чем выше частота волны, тем более высокая частота сигнала. Косинусоидальные волны высокой частоты (цвет на изображении меняется резко) влияют на изображение в меньшей степени, чем волны низкой частоты и в процессе квантования принимают значения близкие к нулю и в процессе кодирования отсекаются.

В данном методе сжатия применяются коды переменной длины. Короткие коды присваиваются часто встречающимся пикселям, а длинные – редко встречающимся [10]. При кодировании видеопотока видекодек сопоставляет отдельно взятый пиксел с количеством бит, благодаря этому и происходит распаковка и сжатие информации. В процессе квантования коэффициентов можно выявить, какие коэффициенты в наименьшей мере влияют на изображение и отбросить их в процессе сжатия. В процессе сжатия необходим ДКП матрицы. Ниже представлены формулы для расчёта ДКП матрицы $N \times N$, строки которой - функции косинуса

$$ДКП(m, n) = \sqrt{\frac{1 - \Delta(m,1)}{N} \cos \frac{\pi(n - 1/2)(m - 1)}{N}}$$

где

$$\Delta(m,1) = \begin{cases} 1, & \text{если } m = 1 \\ 0, & \text{в противном случае} \end{cases}$$

N - число выборок в блоке [11].

Данный метод основывается на прохождении сигнала через фильтры верхних (ФВЧ) и нижних (ФНЧ) частот. Перед тем, как сигнал изображения будет передан на фильтры, происходит его разбиение на две равнозначных группы. Наиболее важным условием сжатия является воспроизведение первоначального изображения с минимальными потерями, это необходимо учитывать при подборе параметров фильтров. Выполнение данного условия позволит подать результат на другую пару фильтров для большего сжатия [12].

Простым примером DWT является вейвлет – преобразование Хаара, которое выполняется в следующем порядке.

1. Входной сигнал - $x[n]$ (ряд выборок, индексированных по n).
2. Нахождение среднего числа двух последовательных выборок.

$$G[n] = \frac{1}{2} (x[n] + x[n + 1])$$

3. Нахождения разности двух последовательных выборок.

$$H[n] = \frac{1}{2} (x[n + 1] - x[n])$$

Заметим, что:

$$X[n] + g[n] = h[n]$$

$$X[n + 1] = g[n] + h[n]$$

В результате применения данного метода кодирования на выходе возможно получить избыточную информацию, благодаря этому, опустив четные члены последовательностей

сигналов, возможно получить две последовательности. Далее воспроизводится входной сигнал $x[n]$.

$$g[0], g[2], g[4], \dots$$

$$h[0], h[2], h[4], \dots$$

$$X[0] = g[0] - h[0]$$

$$X[1] = g[0] + h[0]$$

$$X[2] = g[2] - h[2]$$

$$X[3] = g[2] + h[2]$$

и т.д.

Отличительной особенностью ФНЧ является то, что если на его вход подать сигнал изображения, то разрешение выходного изображения будет значительно ухудшаться. Чтобы решить данную проблему и частично привести качество полученного изображения к исходному виду применяется фильтр высоких частот, который добавляет детали к сжатому изображению [13].

Данный метод сжатия выполняется в два этапа. Первый этап – это основной уровень, когда изображение подается на ФНЧ, а второй – это уровень расширения, на котором происходит процесс добавления деталей в изображение для повышения его разрешения.

Данный метод можно применять несколько раз к одному и тому же видео файлу для большего сжатия.

Так как с каждым днем повышается разрешение видеофайлов, насыщенность цвета на экране становится все больше, растут и требования к кодекам. Необходимо применять все более сложные фильтры. Из-за примитивности данных фильтров, такой метод сжатия в настоящее время на практике не применяется [14].

Метод компенсации движения

Как сказано выше, видеоизображение – это последовательность кадров. В большинстве современных видеоигр частота величина частоты кадров в секунду равна 60, а для видеофайлов – 30. Если сравнить два соседних кадра, то они не будут иметь существенных различий, данный алгоритм основан на поиске сходств между кадрами, которые следуют друг за другом. Необходимо отметить, что проще всего сжимать видео, которые сняты одним кадром – «без склеек», так как в данном случае отсутствует резкая смена кадра и сходство между соседними кадрами максимальное. Данный алгоритм разбивает кадр на блоки, выполняется поиск аналогичного блока в другом кадре (оценка движения), и на основе этих данных создается поле векторов движения. При выполнении компенсации необходимо учитывать вектора движения, это позволит создать изображение максимально близкое к исходному (Рисунок 3).

Подробнее рассмотрим этапы данного метода кодирования.

1. На первом этапе на вход подается текущий кадр видеопотока;
2. Кадр делится на ячейки одного размера (например 8x8 пикселей);
3. Далее происходит обработка каждого блока, в ходе которой выполняется поиск максимальных соответствий того же самого блока на предыдущем кадре.
4. Следующим шагом возможно получить отображение разницы нахождения блоков на текущем кадре, относительно предыдущего.

5. Заключительным этапом создается изображение на основе разницы соседних кадров (Рисунок 4) [5].

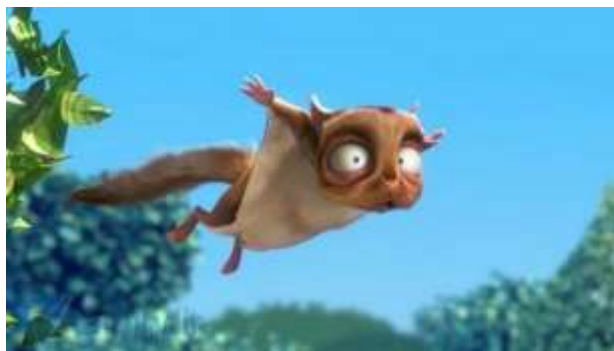


Рисунок 3 – Исходное изображение

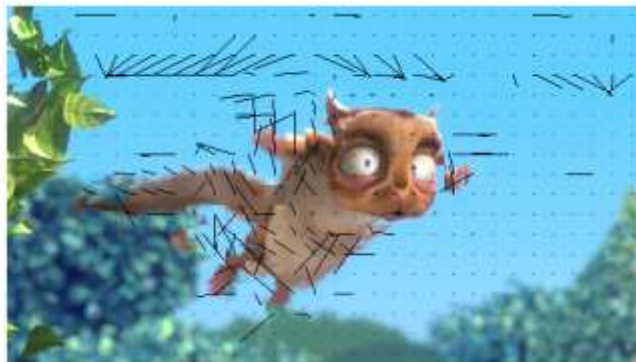


Рисунок 4 – Сжатый кадр

Заключение

Из сравнения рассмотренных методов сжатия следует, что наиболее удобным методами является методы дискретного косинусного преобразования и векторной квантизации, данные методы подразумевают деление изображения на блоки, квантование сигнала и группировку блоков по общим признакам. Данные методы сжатия могут применяться для полноцветных изображений и не зависят от изменения яркости, тогда как методы кодирования длин серий и дискретного косинусного преобразования не применяются для сжатия полноцветных изображений, а метод компенсации движения чувствителен к изменению яркости изображения.

Таким образом, кодирование с потерями обеспечивает сжатие видеоданных, но при этом качество декодированного видео снижается. В результате сравнения основных видов сжатия изображений и основных принципов их работы, пришли к выводу, что в основном процесс кодирования построен на разделении исходного файла на блоки, для последующей их группировки, что дает значительное уменьшение кодированного файла по отношению к первоначальному.

Список литературы

1. Крючкова, Ю. П. Анализ искажений изображений в современных видеокодеках / Ю. П. Крючкова, Ф. М. Игнатов, И. В. Власюк // Фундаментальные проблемы радиоэлектронного приборостроения. – 2009. – Т. 9, № 4. – С. 182-185.
2. Лапшенков, Е. М. Модель оценки потерь качества графического изображения при сжатии с потерями, ориентированная на системы распознавания образов / Е. М. Лапшенков // Компьютерная оптика. – 2011. – Т. 35, № 3. – С. 408-415.
3. Филиппов, Т. К. Сжатие цифровых данных с потерями. Оценка качества сжатия / Т. К. Филиппов // Северный регион: наука, образование, культура. – 2013. – № 2(28). – С. 100-105.
4. Судаков, С. Е. Протоколы сжатия видеоинформации в Internet / С. Е. Судаков // Актуальные проблемы радио- и кинотехнологий : материалы II Международной научно-технической конференции, Санкт-Петербург, 24–27 октября 2017 года. – Санкт-Петербург: Санкт-Петербургский государственный институт кино и телевидения, 2018. – С. 130-136.
5. Исследование методов сегментации изображений / И. В. Грибков, А. В. Захаров, П. П. Кольцов [и др.] // Программные продукты и системы. – 2008. – № 4. – С. 7.
6. Станкевич, Л. А. Распознавание трехмерных объектов с использованием их структурного описания / Л. А. Станкевич, В. В. Тихомиров, Д. В. Троцкий // Нейрокомпьютеры: разработка, применение. – 2006. – № 6. – С. 70-79.
7. Токтасынов, С. Э. Анализ программ сжатия видео с выявлением качества работы кодеков / С. Э. Токтасынов // Системы управления, информационные технологии и математическое моделирование : Материалы V Всероссийской научно-практической конференции с международным участием, Омск, 25–26 апреля 2023 года / Отв. редактор В.А. Бадрызлов. – Омск: Омский государственный технический университет, 2023. – С. 521-528.
8. Акимов, В. А. Дистанционные технологии в образовании. Алгоритмы сжатия информации и форматы данных для передачи текстовой, звуковой и видеоинформации / В. А. Акимов // Известия МГТУ МАМИ. – 2013. – Т. 2, № 4(18). – С. 352-355.
9. Донец, Е. С. Анализ методов кодирования видеоинформации для условий передачи видео в канале с малыми скоростями / Е. С. Донец // Актуальные тенденции и инновации в развитии российской науки : Сборник научных статей. Том Ч. XI. – Москва : Издательство "Перо", 2021. – С. 61-68.
10. Гущина, О. Н. Удаление артефактов блочности из сжатых изображений с использованием адаптивного дискретного косинусного преобразования / О. Н. Гущина, В. Е. Соловьев, А. Н. Ганин // DSPA: Вопросы применения цифровой обработки сигналов. – 2011. – Т. 1, № 3. – С. 119-122.
11. Токарь, П. С. Особенности применения дискретного косинусного преобразования в алгоритмах сжатия данных с потерями / П. С. Токарь, Е. Н. Шевченко // Вестник современных исследований. – 2018. – № 1.1(16). – С. 139-141.
12. Кобер, В. И. Рекурсивное вычисление дискретного косинусного преобразования для обработки сигналов с пониженным разрешением / В. И. Кобер, А. В. Вохминцев, А. В. Мельников // Доклады Академии наук. – 2015. – Т. 461, № 3. – С. 257.

13. Евсютин, О. О. Встраивание информации в область дискретного косинусного преобразования цифровых изображений с безошибочным извлечением / О. О. Евсютин, А. С. Кокурина // Электронные средства и системы управления. Материалы докладов Международной научно-практической конференции. – 2017. – № 1-2. – С. 70-73.
14. Веденин, А. В. Реализация дискретного косинусного преобразования для сжатия изображений и видео с применением вычислений на графических процессорах / А. В. Веденин, С. М. Морозов // Наука настоящего и будущего. – 2018. – Т. 1. – С. 89-90.
15. Потапов, П. В. Использование методов локальной оптимизации для поиска векторов похожести при сжатии видеоданных / П. В. Потапов // Электронные средства и системы управления. Материалы докладов Международной научно-практической конференции. – 2009. – № 1. – С. 60-66.

References

1. Kryuchkova, Yu. P. Analysis of image distortions in modern video codecs / Yu. P. Kryuchkova, F. M. Ignatov, I. V. Vlasyuk // Fundamental problems of radioelectronic instrumentation. – 2009. – Vol. 9, No. 4. – pp. 182-185.
2. Lapshenkov, E. M. A model for estimating the loss of graphic image quality during lossy compression, focused on image recognition systems / E. M. Lapshenkov // Computer optics. - 2011. – vol. 35, No. 3. – pp. 408-415.
3. Filippov, because Lossy digital data compression. Assessment of compression quality / T. K. Filippov // Northern region: science, education, culture. – 2013. – № 2(28). – Pp. 100-105.
4. Sudakov, S. E. Protocols for compressing video information on the Internet / S. E. Sudakov // Actual problems of radio and film technologies : materials of the II International Scientific and Technical Conference, St. Petersburg, October 24-27, 2017. – St. Petersburg: St. Petersburg State Institute of Film and Television, 2018. – pp. 130-136.
5. Investigation of image segmentation methods / I. V. Gribkov, A.V. Zakharov, P. P. Koltsov [et al.] // Software products and systems. - 2008. – No. 4. – S. 7.
6. Stankevich, L. A. Recognition of three-dimensional objects using their structural description / L. A. Stankevich, V. V. Tikhomirov, D. V. Trotsky // Neurocomputers: development, application. - 2006. – No. 6. – pp. 70-79.
7. Toktasynov, S. E. Analysis of video compression programs with identification of the quality of codecs / S. E. Toktasynov // Control systems, information technologies and mathematical modeling : Materials of the V All-Russian Scientific and practical conference with international participation, Omsk, April 25-26, 2023 / Editor V.A. Badryzlov. Omsk: Omsk State Technical University, 2023. pp. 521-528.
8. Akimov, V. A. Distance learning technologies in education. Information compression algorithms and data formats for the transmission of text, audio and video information / V. A. Akimov // Izvestiya MSTU MAMI. – 2013. – Vol. 2, No. 4(18). – pp. 352-355.
9. Donets, E. S. Analysis of video encoding methods for video transmission conditions in a channel with low speeds / E. S. Donets // Current trends and innovations in the development of Russian science : A collection of scientific articles. Volume XI. – Moscow : Publishing House "Pero", 2021. – pp. 61-68.

10. Gushchina, O. N. Removal of blockiness artifacts from compressed images using adaptive discrete cosine transformation / O. N. Gushchina, V. E. Solovyov, A. N. Ganin // DSPA: Issues of digital signal processing application. – 2011. – Vol. 1, No. 3. – pp. 119-122.
 11. Tokar, P. S. Features of the application of discrete cosine transformation in lossy data compression algorithms / P. S. Tokar, E. N. Shevchenko // Bulletin of Modern Research. – 2018. – № 1.1(16). – Pp. 139-141.
 12. Kober, V. I. Recursive calculation of a discrete cosine transform for low-resolution signal processing / V. I. Kober, A.V. Vokhmintsev, A.V. Melnikov // Reports of the Academy of Sciences. – 2015. – Vol. 461, No. 3. – p. 257.
 13. Evsyutin, O. O. Embedding information in the field of discrete cosine transformation of digital images with error-free extraction / O. O. Evsyutin, A. S. Kokurina // Electronic means and control systems. Materials of the reports of the International Scientific and Practical Conference. - 2017. – No. 1-2. – pp. 70-73.
 14. Vedenin, A.V. Implementation of discrete cosine transformation for image and video compression using calculations on graphics processors / A.V. Vedenin, S. M. Morozov // Science of the present and the future. – 2018. – Vol. 1. – pp. 89-90.
 15. Potapov, P. V. Using local optimization methods to search for vectors similarities in video data compression / P. V. Potapov // Electronic tools and control systems. Materials of the reports of the International Scientific and Practical Conference. - 2009. – No. 1. – pp. 60-66.
-