

## АЛГОРИТМЫ СЖАТИЯ АУДИОСИГНАЛОВ БЕЗ ПОТЕР И С ПОТЕРЯМИ

*Ш.Р. Юсупов – преподаватель кафедры  
«Компьютерные науки» Ургенчского  
государственного университета*

• В статье анализированы алгоритмы сжатия аудиосигналов без потер и с потерями.

• Ушбу мақолада аудиосигналларни йўқотишли зичлаш ва йўқотишсиз зичлаш алгоритмлари таҳлил қилинган.

• The article analyzes audio compression algorithms without lossy and lossy.

**Ключевые слова:** Звук, дискретизация, квантование, цифровой звук, метод Лемпеля–Зива–Уэлча, методе Берроуза–Уиллера, алгоритме Хаффмана, подавления пауз, Аудиофайл, сэмпл, Формат MPEG Layer-3, алгоритм сжатия MPEG, преобразования Фурье, обратное преобразование

**Калит сўзлар:** товуш, дискретлаш, квантлаш, рақамли товуш, Лемпел – Змва – Уэлч методи, Берроуз – Уиллер методи, Хаффман алгоритми, тўхталишларни зичлаш, аудиофайл, сэмпл, MPEG Layer-3 формати, MPEG зичлаш алгоритми, Фурье алмаштириши, тескари алмаштириш.

**Key words:** Sound, sampling, quantization, digital sound, Lempel-Ziva-Welch method, Burroughs-Wheeler method, Huffman algorithm, pause suppression, Audio file, sample MPEG Layer-3 format, MPEG compression algorithm, Fourier transforms, inverse transform.

Звук – это упругое колебание среды, он распространяется в среде с помощью волн давления посредством колебания атомов и молекул. Как и любая волна, звук характеризуется скоростью, амплитудой и частотой.

С записью звука в цифровом виде связано два действия: дискретизация и квантование.

Дискретизация. В соответствии с теоремой отсчетов В.А. Котельникова неискаженная передача непрерывного (аналогового) сигнала с полосой частот  $0..F_{\max}$  дискретной последовательностью его отсчетов возможна, если частота дискретизации  $f_d$  связана с максимальной частотой  $F_{\max}$  исходного сигнала соотношением:

$$f_d > 2F_{\max}. \quad (1)$$

Квантование. Квантование – это преобразование аналогового сигнала в

ступенчатый сигнал с двоичным отсчетом уровней в квантах. При этой операции производится округление входного сигнала к принятой двоичной шкале квантователя. Процедуру квантования можно рассматривать как прохождение входного сигнала через устройство с амплитудной характеристикой ступенчатой формы, которая называется характеристикой (или шкалой) квантования. Если в пределах всей шкалы шаг квантования постоянен ( $x_i - x_{i-1} = \Delta$  и  $y_i - y_{i-1} = \Delta$ ), то квантование называют равномерным. Такой вид квантования удобен для начального представления звукового сигнала с целью последующей обработки.

Таким образом, главными параметрами цифрового звука (аудиосигнала) является частота дискретизации (ЧД) и количество ступеней амплитуды (обозначаемое двоичной разрядностью отсчета - РО). ЧД определяет частотный диапазон - прозрачность, звонкость звука, а РО - количество шумов, или "гладкость" звука.

Человеческое ухо воспринимает в большинстве случаев звук с частотой не выше 22000 Гц, и для того чтобы его полностью описать в цифровом виде, по формуле (1) требуется частота дискретизации не менее 44,1 кГц. Так как абсолютно точно определить значение сигнала в определенный момент времени невозможно, то при оцифровке происходит квантование, то есть замена реальных значений сигнала приближенными. Чем больше уровней квантования звука, тем точнее описывается уровень сигнала. В итоге каждый стандартный компакт-диск несет на себе звуковой сигнал с частотой дискретизации в те самые 44,1 кГц и уровнем квантования в 16 бит, а в некоторых устройствах производится дискретизация с частотой 48 кГц.

К настоящему времени разработано множество алгоритмов кодирования общего назначения, которые выискивают повторяющиеся последовательности в двоичных данных. Если им удастся обнаружить такие последовательности, то они могут сжать данные, заменив последовательности кодами. Модуль, производящий декомпрессию (которому доступна информация о таких последовательностях) может произвести обратную операцию. Эти алгоритмы относятся к алгоритмам сжатия без потерь, так как восстановленные данные будут с точностью до бита идентичны исходным. Они могут быть использованы для компрессии любых данных, но при сжатии аудио сигналов оказываются мало эффективными.

Например, в методе Лемпеля–Зива–Уэлча (Lempel–Ziv–Welch, LZW), дефляции и методе Берроуза–Уиллера (Burroughs–Wheeler) алгоритм заключается в поиске длинных последовательностей байтов, встречающихся в файле несколько раз. Однако, в звуковых файлах длинных повторяющихся последовательностей обычно не много: в основном это обуславливается наличием шумов.

В алгоритме Хаффмана (Huffman) и арифметическом кодировании ищутся байты с определенными значениями (или пары байтов), попадающие в файле чаще остальных. Как только удастся выделить такое значение, строится код, который тем короче, чем чаще встречается значение. Действительно, у звуковых файлов неравномерно распределены возможные значения байтов, поэтому алгоритмы данного типа могут достаточно хорошо обрабатывать звуковые файлы. Но так как звуковые файлы могут обладать очень большим объемом, эти методы также становятся неэффективными. Как правило, методы сжатия без потерь используются в современных стандартах сжатия на вторичном, дополнительном этапе сжатия с целью повышения коэффициента сжатия.

Еще одним недостатком алгоритмов сжатия без потерь является неоднородность компрессии, что не позволяет наложить ограничения на скорость передачи сжатых аудио потоков.

Можно добиться лучших результатов при сжатии звука с потерей части аудиоинформации, развивая методы компрессии, которые учитывают особенности восприятия звука. Они удаляют ту часть данных, которая остается неслышимой для органов слуха. Это похоже на сжатие изображений с отбрасыванием информации, незаметной для глаза. В обоих случаях мы исходим из того факта, что исходная информация (изображение или звук) является аналоговым, то есть, часть информации уже потеряно при квантовании и оцифровывании. Если допустить еще некоторую потерю, сделав это аккуратно, то это не повлияет на качество воспроизведения разжатого звука, который не будет сильно отличаться от оригинала. Мы кратко опишем подходы которые называются подавлением пауз и уплотнением.

Идея подавления пауз заключается в рассмотрении малых сэмплов, как если бы их не было (то есть, они равны нулю). Такое обнуление будет порождать серии нулей, поэтому метод подавления пауз, на самом деле, является вариантом RLE, приспособленным к сжатию звука. Этот метод основан на особенности звукового восприятия, которое состоит в терпимости уха человека к отбрасыванию еле слышных звуков. Аудиофайлы, содержащие длинные участки тихого звука будут лучше сжиматься методом подавления пауз, чем файлы, наполненные громкими звуками. Этот метод требует участие пользователя, который будет контролировать параметры, задающие порог громкости для сэмплов. При этом необходимы еще два параметра, они не обязательно контролируются пользователем. Один параметр служит для определения самых коротких последовательностей тихих сэмплов, обычно, это 2 или 3. А второй задает наименьшее число последовательных громких сэмплов, при появлении которых прекращается тишина или пауза. Например, после 15 таких сэмплов может последовать 2 громких, а затем 13 тихих, что будет определено как одна

большая пауза длины 30, а аналогичная последовательность из 15, 3 и 12 сэмплов, станет двумя паузами с коротким звуком между ними.

Формат MPEG Layer-3 (сокращенно MP3). При кодировании в формат файлов MPEG Layer-3 исходный звуковой файл режется на фрагменты, длительностью по 50 миллисекунд, каждый из которых анализируется отдельно.

При анализе фрагмент раскладывается на гармоники по методу Фурье, из которых в соответствии с теорией восприятия звука человеческим ухом выбрасываются те гармоники, которые человек не воспринимает или воспринимает хуже на фоне остальных. Кроме того, выбрасываются звуки, замаскированные вследствие инертности слуха. Информация об оставшихся после фильтрации гармониках и записывается в MP3 файл, который в результате получается гораздо меньше по размеру, чем исходный WAV.

При воспроизведении производится обратное преобразование, при котором оставшиеся гармоники вновь преобразуются в звуковую волну. Получающийся звук не совпадает с исходным, но поскольку отбрасывались малозначимые и не слышимые звуки, отличить сигнал от исходного для человеческого уха достаточно сложно.

Алгоритм сжатия MPEG можно разделить на три этапа:

- предварительная обработка;
- основное преобразование;
- кодирование и упаковка компонент преобразования.

На этапе предварительной обработки производится, в общем случае, подготовка исходного потока аудио данных к выполнению процедуры основного преобразования. В частности, можно выделить два вида такой подготовки: разбиение на блоки и фильтрация шумов.

На втором этапе с помощью дискретного преобразования Фурье (ДПФ) входные отсчеты ИКМ преобразуются в 512 спектральных составляющих. Таким образом осуществляется переход от временного представления сигнала к частотному.

На этапе кодирования и упаковки компонент производится анализ частотной области психоакустической моделью, которая отбрасывает неслышимые компоненты спектра и вычисляет шаг квантования, при котором шум квантования будет не слышен. Также на этом этапе осуществляется само квантование оставшихся спектральных отсчетов и, далее, они подвергаются кодированию по методу Хаффмана.

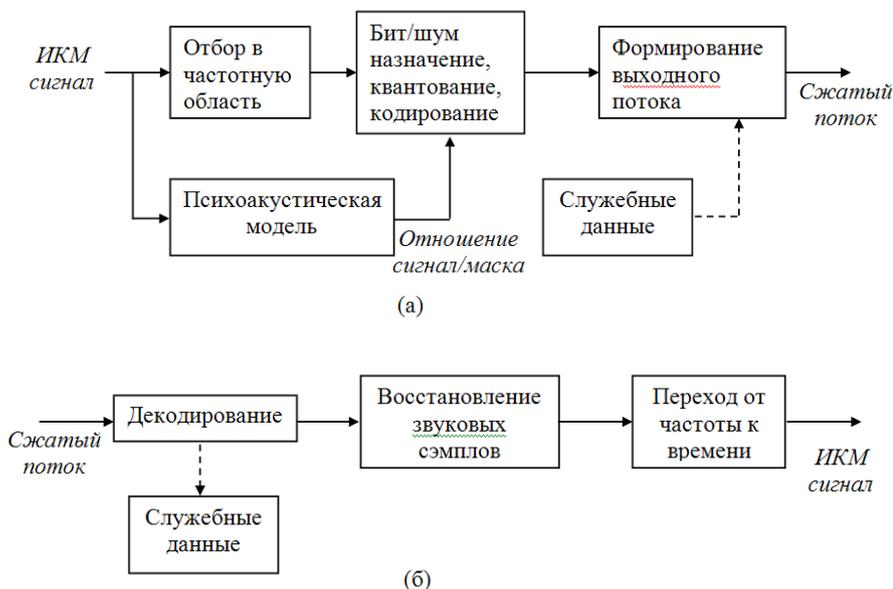


Рис.1. Кодер (а) и декодер (б) MPEG.

Алгоритм восстановления сигнала состоит из двух этапов:

- восстановление отсчетов;
- обратное преобразование.

На этапе восстановления отсчетов происходит декодирование спектральных компонент. На втором этапе с помощью обратного ДПФ производится переход к временному представлению сигнала. На рисунке приведены структурные схемы кодера и декодера MPEG.

### Литература:

1. А.Ю. Тропченко, А.А. Тропченко. Методы сжатия изображений, аудиосигналов и видео. Санкт-Петербург 2009 г.
2. Дэвид Сэломон. Сжатие данных, изображений и звука. Техносфера–2006 г.
3. Юкио Сато. Обработка сигналов, первое знакомство. Додэка 2002 г.