

Кодирование звука с потерями

Сжатие данных с потерями

- Сжатие данных с потерями — метод сжатия (компрессии) данных, при использовании которого распакованные данные отличаются от исходных, но степень отличия не является существенной с точки зрения их дальнейшего использования.
- Применяются для сжатия:
 - Аудиоданных;
 - Видеоданных;
 - Статических изображений;
 - При потоковой передаче данных;
 - В цифровой телефонии.

Достоинства и недостатки

- Достоинство:
 - Степень сжатия;
- Недостатки:
 - Потеря информации;
 - повторное сжатие с потерями снижает качество, а декодирование увеличивает размер, не возвращая или не повышая качество.

Битрейт

- Битрейт (англ. bit rate) — буквально, скорость прохождения битов информации.
- В форматах потокового видео и аудио (MPЕG, MP3), использующих сжатие с потерей качества, параметр «битрейт» выражает степень сжатия потока и, тем самым, определяет размер канала, для которого сжат поток данных.
- Чаще всего битрейт звука и видео измеряют в килобитах в секунду (англ. kilobit per second, kbps), реже — в мегабитах в секунду (только для видео).

Типы битрейтов

- **Постоянный битрейт (CBR - Constant Bit Rate);**
- **Изменяющийся битрейт или переменный битрейт (VBR - Variable Bit Rate);**
- **Усредненный битрейт (ABR - Average Bit Rate)**

Постоянный битрейт

- **Постоянный битрейт.** Битрейт не меняется на протяжении всего файла. Каждой секунде произведения соответствует одинаковое количество закодированных бит данных (даже при кодировании тишины).
- **Достоинства:**
 - возможность довольно точно предсказать размер конечного файла;
 - может быть использован для передачи **потоковых** мультимедиа данных по ограниченному каналу связи.
- **Недостаток:** не очень подходит для музыкальных произведений, звучание которых динамично изменяется во времени, так как не обеспечивает оптимального соотношения размер/качество.

Переменный битрейт

- **Переменный битрейт.** Битрейт динамически изменяется программой-кодером при кодировании в зависимости от насыщенности кодируемого аудиоматериала и установленного пользователем качества кодирования (например, тишина закодируется с минимальным битрейтом).
- **Достоинства:**
 - аудиоматериал разной насыщенности может быть закодирован с определённым качеством, которое обычно выше, чем при установке среднего значения в методе CBR;
 - размер файла уменьшается за счёт фрагментов, не требующих высокого битрейта.
- **Недостаток:** сложность предсказания размера выходного файла.

Усреднённый битрейт

- **Усреднённый битрейт** является гибридом постоянного и переменного битрейтов: битрейт в кбит/с задаётся пользователем (как при CBR), а программа варьирует его в некоторых пределах (как при VBR).

Битрейт для аудио CD

- Скорость передачи в битах при импульсно-кодовой модуляции (PCM) аудиоданных может быть рассчитана по следующей формуле:

$$\text{bit rate} = \text{sample rate} \times \text{bit depth} \times \text{channels}$$

- Для стандартного аудио CD:

$$44,100 \times 16 \times 2 = 1,411,200 \text{ bit/s} = 1,411.2 \text{ kbit/s}$$

- Общий размер аудиоданных (исключая размер заголовков и метаданные) рассчитывается как (в битах):

$$\text{size in bits} = \text{sample rate} \times \text{bit depth} \times \text{channels} \times \text{length of time}$$

- 80 минут (4800 секунд) данных на аудио CD требует
- 846 720 000 байт памяти:

$$\frac{44,100 \times 16 \times 2 \times 4,800}{8} = 846,720,000 \text{ bytes} \approx 847 \text{ MB}$$

Примеры битрейтов для аудио

- 800 – 1400 бит/с - минимум, необходимый для узнаваемой речи
- 8 кбит/с - качество телефона (используются речевые кодеки)
- 32-500 кбит/с - аудиос потерями, используемый в Ogg Vorbis
- 256 кбит/с - Цифровое радиовещание
- 400-411 кбит/с - аудио без потерь, используемый в таких форматах, как Free Lossless Audio Codec, WavPack и т.п.
- 1 411.2 кбит/с - звук формата CD-DA.
- 5 644.8 кбит/с - Super Audio CD
- 6,144 Мбит/с - Dolby Digital Plus (на основе AC-3 кодека)
- 18 Мбит/с - расширенный Lossless Audio Codec

Примеры битрейта для видео

- 16 кбит/с - качество видеотелефонии (используются различные схемы сжатия видео для передачи изображения одной головы с речью)
- 128-384 кбит/с - качество видеоконференций с использованием сжатия видео
- 1.15 Мбит/с - максимальное качество VCD (MPEG1)
- 3.5 Мбит/с - качество стандартного телевидения (MPEG-2)
- 9.8 Мбит/с - DVD (MPEG2)
- 8-15 Мбит/с - качество HDTV (MPEG-4 AVC)
- 19 Мбит/с - HDV 720p (MPEG2)
- 25 Мбит/с - HDV 1080i (MPEG2)
- 40 Мбит/с - Blu-Ray Disc (MPEG2, AVC, VC-1)

Основные идеи используемые при сжатии с потерями

- Пренебрежение тонкими деталями звучания оригинала, лежащие вне пределов которые воспринимает человеческое ухо. При этом может использоваться:
 - **Эффект маскировки** (один звук может быть скрыт другим звуком. например, звук сирены может заглушать речь человека);
 - **Деление полосы звуковых частот на подполосы.** Каждая полоса обрабатывается отдельно. Кодировщик выделяет самые громкие звуки в каждой полосе и использует эту информацию для определения приемлемого уровня шума для этой полосы.
 - **Использование психоакустической модели.** Сжатие с использованием этой модели основано на удалении заведомо неслышимых частот с более тщательным сохранением звуков, хорошо различаемых человеческим ухом.
 - **Совмещение стерео.** Известно, что слуховой аппарат человека может определить направление лишь средних частот - высокие и низкие звучат как бы отдельно от источника. Значит, эти фоновые частоты можно кодировать в моно сигнал.

Эффект маскировки

- **Одновременная маскировка.** Более громкий звук снижает восприятие более слабого, вплоть до исчезновения его слышимости. Чем ближе частота маскируемого звука к частоте маскирующего, тем сильнее он будет скрываться.
- **Временная маскировка.** При прекращении подачи маскирующего звука маскируемый некоторое время продолжает быть неслышимым.
- **Постстимульное утомление.** Нередко после воздействия громких звуков высокой интенсивности у человека резко снижается слуховая чувствительность.

Первый тип сжатия с потерями

- В трансформирующих кодеках фреймы изображений или звука обычно трансформируются в новое базисное пространство и производится квантование.
- Трансформация может осуществляться:
 - для всего фрейма целиком (как, например, в схемах на основе wavelet-преобразования)
 - поблочно (JPEG, MP3)
- Результат затем сжимается энтропийными методами.

Второй тип сжатия с потерями

- В предсказывающих кодеках предыдущие и/или последующие отсчеты данных используются для того, чтобы предсказать текущий отсчет изображения или звука.
- Ошибка между предсказанными данными и реальными вместе с добавочной информацией, необходимой для производства предсказания, затем квантуется и кодируется.

MP3

- MP3 - самых распространённых и популярных форматов цифрового кодирования звуковой информации с потерями.
- Принцип сжатия заключается в снижении точности некоторых частей звукового потока, что практически неразлично для слуха большинства людей. Данный метод называют **кодированием восприятия**.
- На первом этапе строится диаграмма звука в виде последовательности коротких промежутков времени (блоков), затем на ней удаляется информация, не различимая человеческим ухом, а оставшаяся информация сохраняется в компактном виде в свой **фрейм** (кадр).

Битрейт для MP3

- 32 кбит/с - в целом приемлемыми только для речи
- 96 кбит/с - обычно используется для речи или некачественного потокового аудио
- 128 или 160 кбит/с – битрейт для среднего качества
- 192 кбит/с - обычно используется битрейт для высококачественного аудио
- 320 кбит/с – наивысший уровень битрейта, поддерживаемый в MP3

Кодирование стерео в MP3

- **Dual Channel** - Каждый канал получает ровно половину потока и кодируется отдельно как моно сигнал.
- **Stereo** - Каждый канал кодируется отдельно, но кодер может принять решение отдать одному каналу больше места, чем другому.
- **Joint Stereo (MS Stereo)** - Стереосигнал раскладывается на средний между каналами и разностный. При этом второй кодируется с меньшим битрейтом.

Кодеки в MP3

- MP3 является контейнером. **Контейнер** - это формат файла, определяющей распределение аудио, видео информации внутри него. Типом контейнера в большинстве случаев не выбирается тип кодирования (сжатия) информации внутри файла.
- В MP3 могут использоваться следующие кодеки:
 - LAME
 - MP3 Compressor
 - Blade's MP3 Encoder
 - MP3 Encoder
 - mpegEnc
 - XingMP3 Encoder
 - AudioCatalyst
- <http://www.ixbt.com/multimedia/mp3-codec-tst1.html>