

Audio Format Comparison

MP3 Audio

MP3 (MPEG-1 Audio Layer III) — это аудиокодек, разработанный для сжатия аудиоданных с потерями, который стал одним из самых популярных форматов для хранения и передачи музыки в цифровом виде. MP3 был разработан как часть проекта MPEG-1 и позже расширен в MPEG-2. Этот кодек использует различные алгоритмы и методы для уменьшения размера файла при сохранении высокого качества звука. Ниже приведено подробное описание алгоритмов, методов и функций, которые используются в MP3.

Основные алгоритмы и методы в MP3

1. Психоакустическая модель (Psychoacoustic Model)

- **Маскирование:** MP3 использует эффекты маскирования, чтобы избавиться от звуков, которые не слышны человеческим ухом. Маскирование бывает двух типов:
 - **Спектральное маскирование:** Громкие звуки на определенных частотах могут маскировать более тихие звуки на соседних частотах.
 - **Временное маскирование:** Громкий звук может маскировать более тихие звуки, которые происходят незадолго до или после него.
- Психоакустическая модель анализирует аудиосигнал и определяет, какие части спектра можно удалить или уменьшить по громкости без заметного ухудшения качества.

2. Разделение на фреймы и суббэнды (Frame and Subband Division)

- Аудиосигнал делится на небольшие участки, называемые фреймами, которые обычно имеют длительность 26 миллисекунд.
- Каждый фрейм дополнительно делится на 32 суббэнда с помощью фильтра полифазного анализа (polyphase filter bank), что позволяет анализировать аудиосигнал по частотным компонентам.

3. Быстрое преобразование Фурье (Fast Fourier Transform, FFT)

- MP3 использует FFT для преобразования временного сигнала в частотную область. Это позволяет анализировать и кодировать сигнал по частотным компонентам.

4. Дискретное косинусное преобразование (Modified Discrete Cosine Transform, MDCT)

- После анализа суббэндов, каждый суббэнд подвергается MDCT , что позволяет более эффективно представлять данные в частотной области и улучшает сжатие.

5. Квантование и кодирование (Quantization and Coding)

- **Квантование:** Коэффициенты, полученные после MDCT , квантируются. Это означает, что они преобразуются в дискретные значения с меньшим числом бит, что снижает точность, но уменьшает объем данных.
- **Huffman Coding:** После квантования данные подвергаются энтропийному кодированию с использованием кодирования Хаффмана, что позволяет еще больше уменьшить размер данных без потерь.

6. Битрейт и распределение битов (Bitrate and Bit Allocation)

- MP3 поддерживает различные битрейты, что позволяет контролировать соотношение качества и размера файла.
- Кодек динамически распределяет биты между фреймами и суббэндами на основе психоакустической модели и сложности сигнала, чтобы оптимизировать качество звука.

7. Постобработка (Post-processing)

- **Антиалиасинг (Anti-aliasing):** Применяется для удаления артефактов, которые могут возникнуть при фильтрации.
- **Сtereo кодирование (Stereo Encoding):** MP3 использует несколько методов стерео кодирования для уменьшения объема данных:
 - **Mid/Side (M/S) Coding:** Разделение стереосигнала на средний (суммарный) и боковой (разностный) каналы.
 - **Intensity Stereo Coding:** Использование интенсивности звука для представления стереопанорамы при низких битрейтах.

Пример работы MP3

1. **Разделение на фреймы:** Входной аудиосигнал делится на небольшие фреймы длительностью около 26 миллисекунд.
2. **Анализ психоакустической модели:** Психоакустическая модель определяет, какие звуки можно удалить или уменьшить по громкости.
3. **Фильтрация и разделение на суббэнды:** Аудиосигнал фильтруется и разделяется на 32 суббэнда с помощью полифазного фильтра.
4. **Преобразование в частотную область:** Каждый суббэнд подвергается FFT и MDCT , преобразуя сигнал в частотную область.

5. **Квантование:** Коэффициенты MDCT квантуются для уменьшения объема данных.
6. **Энтропийное кодирование:** Квантованные данные кодируются с использованием кодирования Хаффмана.
7. **Распределение битов:** Биты динамически распределяются между фреймами и суббэндами для оптимизации качества.
8. **Постобработка:** Применяются антиалиасинг и стерео кодирование для улучшения качества звука.
9. **Сборка фреймов:** Квантованные и закодированные данные собираются в выходные MP3 фреймы.

Заключение

MP3 остается популярным форматом аудиокодирования благодаря своей эффективности и высокой совместимости. Он использует множество сложных алгоритмов и методов, таких как психоакустическая модель, фильтрация, преобразование в частотную область, квантование и энтропийное кодирование, чтобы достичь значительного сжатия аудиоданных при минимальных потерях качества. Эти методы позволяют MP3 обеспечивать хорошее качество звука даже при относительно низких битрейтах.

AAC Audio

AAC (Advanced Audio Coding) — это аудиокодек, разработанный для замены MP3 и обеспечения более высокой эффективности сжатия при сохранении или улучшении качества звука. AAC является частью стандартов MPEG-2 и MPEG-4 и широко используется в различных приложениях, включая потоковое аудио, цифровое телевидение и мобильные устройства. Ниже приведено подробное описание алгоритмов, методов и функций, которые используются в AAC.

Основные алгоритмы и методы в AAC

1. Психоакустическая модель (Psychoacoustic Model)

- **Маскирование:** AAC использует эффекты спектрального и временного маскирования для устранения звуков, которые не слышны человеческим ухом. Маскирование помогает определить, какие части спектра можно удалить или уменьшить по громкости без заметного ухудшения качества.
- Психоакустическая модель анализирует аудиосигнал и определяет пороги маскирования для каждого фрейма.

2. Разделение на фреймы и блоки (Frame and Block Division)

- Аудиосигнал делится на фреймы длительностью 1024 или 960 семплов (длинные блоки) или 128 или 120 семплов (короткие блоки) для обеспечения гибкости при кодировании различных типов сигналов.
- Длинные блоки используются для кодирования стационарных сигналов, а короткие блоки — для переходных процессов и импульсных сигналов.

3. Модуляция адаптивного модуляционного преобразования (Modified Discrete Cosine Transform, MDCT)

- Каждый фрейм аудиосигнала подвергается MDCT, что позволяет преобразовать временной сигнал в частотную область и эффективно представлять данные для сжатия.

4. Параллельное фильтрование и PQF (Polyphase Quadrature Filter, PQF)

- Используется для разделения аудиосигнала на более мелкие поддиапазоны частот, что улучшает точность и эффективность кодирования.

5. Тональная и нетональная обработка (Tonality Detection and Processing)

- AAC кодирует тональные и нетональные компоненты сигнала отдельно, что улучшает качество звука.
- Тональные компоненты кодируются более точно, а нетональные компоненты могут быть подвергнуты более агрессивному сжатию.

6. Интенсивное стерео (Intensity Stereo) и стерео M/S (Mid/Side)

- AAC использует различные методы стерео кодирования для уменьшения объема данных:
 - **Intensity Stereo:** Использует интенсивность звука для представления стереопанорамы при низких битрейтах.
 - **Mid/Side (M/S) Stereo:** Разделение стереосигнала на средний (суммарный) и боковой (разностный) каналы для уменьшения корреляции между каналами.

7. Перцептивное кодирование (Perceptual Noise Substitution, PNS)

- Метод, который заменяет шумовые компоненты сигнала сгенерированными шумами на стороне декодера, что позволяет существенно снизить объем данных.

8. Алгоритмы квантования и кодирования (Quantization and Coding Algorithms)

- **Квантование:** Коэффициенты MDCT квантуются с учетом психоакустической модели, что снижает объем данных.
- **Безопасное кодирование:** Используются различные методы энтропийного кодирования, включая:

- **Кодирование Хаффмана:** Для эффективного представления часто встречающихся значений.
- **Arithmetic Coding:** Для более гибкого и эффективного кодирования данных.

9. Адаптивное битовое распределение (Adaptive Bit Allocation)

- Биты динамически распределяются между частотными компонентами в зависимости от их важности для перцептивного качества, что оптимизирует качество звука.

10. Динамическое управление диапазоном (Dynamic Range Control, DRC)

- Используется для управления динамическим диапазоном аудиосигнала, улучшая воспроизведение на устройствах с ограниченными возможностями динамиков и в шумных условиях.

Пример работы ААС

1. **Разделение на фреймы:** Входной аудиосигнал делится на фреймы.
2. **Анализ психоакустической модели:** Определяются пороги маскирования и важность различных частотных компонентов.
3. **Преобразование в частотную область:** Каждый фрейм подвергается MDCT для преобразования временного сигнала в частотную область.
4. **Квантование:** Частотные коэффициенты квантируются с учетом порогов маскирования.
5. **Энтропийное кодирование:** Квантованные данные кодируются с использованием кодирования Хаффмана и арифметического кодирования.
6. **Распределение битов:** Биты распределяются между частотными компонентами для оптимизации качества звука.
7. **Стерео кодирование:** Применяются методы интенсивного стерео и M/S стерео для уменьшения объема данных.
8. **Перцептивное кодирование:** Применяется перцептивное кодирование для замены шумовых компонентов сгенерированными шумами.
9. **Сборка фреймов:** Квантованные и закодированные данные собираются в выходные ААС фреймы.

Заключение

ААС является современным и эффективным аудиокодеком, который значительно улучшает качество звука при уменьшении битрейта по сравнению с предыдущими стандартами, такими как MP3. Это достигается за счет использования множества

сложных алгоритмов и методов, включая психоакустическую модель, адаптивное преобразование и квантование, эффективное стерео кодирование, энтропийное кодирование и динамическое управление диапазоном. Эти методы позволяют AAC обеспечивать высокое качество звука даже при низких битрейтах, что делает его идеальным для широкого спектра приложений, от потокового аудио до цифрового телевидения и мобильных устройств.

MP3 VS AAC

MP3 и AAC — это два широко используемых аудиокодека для сжатия звука с потерями. MP3 был разработан в конце 1980-х и начале 1990-х годов и стал стандартом для цифровой музыки. AAC, разработанный позже, предназначен для преодоления ограничений MP3 и обеспечения более высокой эффективности сжатия и лучшего качества звука. Ниже приведено подробное сравнение и основные отличия между MP3 и AAC.

Основные отличия между MP3 и AAC

1. Эффективность сжатия

- MP3 :
 - Разработан как часть стандарта MPEG-1 и MPEG-2.
 - Обеспечивает среднюю степень сжатия с потерями, позволяя уменьшить размер файла примерно до 1/10 от оригинального без сжатия.
 - В MP3 используется фиксированный размер блоков (1152 семплов) для всех типов сигналов.
- AAC :
 - Разработан как часть стандарта MPEG-2 и MPEG-4.
 - Обеспечивает более высокую степень сжатия с потерями по сравнению с MP3, позволяя уменьшить размер файла при сохранении или улучшении качества звука.
 - AAC использует переменный размер блоков (960 или 1024 семплов для длинных блоков и 120 или 128 семплов для коротких блоков), что позволяет более эффективно обрабатывать различные типы сигналов, включая переходные процессы и стационарные звуки.

2. Психоакустическая модель

- MP3 :

- Использует основную психоакустическую модель, которая учитывает эффекты спектрального и временного маскирования для удаления звуков, которые не слышны человеческим ухом.
- Психоакустическая модель MP3 менее точна и эффективна по сравнению с AAC.
- AAC :
 - Использует более усовершенствованную психоакустическую модель, которая обеспечивает более точное маскирование звуков.
 - Учитывает как спектральное, так и временное маскирование, а также другие параметры, что позволяет лучше адаптировать кодирование к особенностям слухового восприятия.

3. Разделение на фреймы и блоки

- MP3 :
 - Разделяет аудиосигнал на фреймы длительностью 26 миллисекунд, каждый из которых содержит 1152 семпла.
 - Использует фиксированный размер блоков для всех типов сигналов, что может быть неэффективно для переходных процессов.
- AAC :
 - Разделяет аудиосигнал на фреймы с переменной длительностью блоков: 1024 или 960 семплов для длинных блоков и 128 или 120 семплов для коротких блоков.
 - Использует длинные блоки для стационарных сигналов и короткие блоки для переходных процессов, что позволяет улучшить качество кодирования.

4. Модуляция адаптивного модуляционного преобразования (MDCT)

- MP3 :
 - Использует модифицированное дискретное косинусное преобразование (MDCT) с фиксированным размером блока.
 - MDCT применяется ко всему фрейму, что может быть неэффективно для переходных сигналов.
- AAC :
 - Использует MDCT с переменным размером блока, что позволяет адаптироваться к различным типам сигналов.

- Применение MDCT с различными размерами блоков улучшает качество кодирования переходных процессов и стационарных звуков.

5. Тональная и нетональная обработка

- MP3 :
 - Не включает специальные методы для отдельной обработки тональных и нетональных компонентов сигнала.
- AAC :
 - AAC кодирует тональные и нетональные компоненты сигнала отдельно, что улучшает качество звука.
 - Тональные компоненты кодируются более точно, а нетональные компоненты могут быть подвергнуты более агрессивному сжатию.

6. Стерео кодирование

- MP3 :
 - Использует методы интенсивного стерео (Intensity Stereo) и совместного стерео (Joint Stereo), включая M/S стерео (Mid/Side) для уменьшения объема данных.
- AAC :
 - AAC включает улучшенные методы стерео кодирования, такие как стерео M/S и интенсивное стерео, что позволяет более эффективно кодировать стереосигналы и уменьшать объем данных.

7. Перцептивное кодирование (PNS)

- MP3 :
 - Не использует перцептивное кодирование шумов.
- AAC :
 - Включает метод перцептивного кодирования шумов (Perceptual Noise Substitution, PNS), который заменяет шумовые компоненты сигнала сгенерированными шумами на стороне декодера, что позволяет существенно снизить объем данных.

8. Алгоритмы квантования и кодирования

- MP3 :

- Использует квантование и кодирование Хаффмана для представления аудиоданных.
- Алгоритмы менее эффективны по сравнению с AAC.
- AAC :
 - Использует усовершенствованные алгоритмы квантования и кодирования, включая арифметическое кодирование и кодирование Хаффмана, что обеспечивает более высокую эффективность сжатия.

9. Динамическое управление диапазоном (DRC)

- MP3 :
 - Не включает встроенное управление динамическим диапазоном.
- AAC :
 - Включает динамическое управление диапазоном (Dynamic Range Control, DRC), которое улучшает воспроизведение на устройствах с ограниченными возможностями динамиков и в шумных условиях.

Заключение

AAC является более современным и эффективным аудиокодеком по сравнению с MP3. Он обеспечивает более высокую степень сжатия и лучшее качество звука при одинаковых битрейтах. Это достигается за счет использования более усовершенствованных методов и алгоритмов, таких как улучшенная психоакустическая модель, переменный размер блоков, адаптивное квантование и кодирование, а также специализированные методы обработки тональных и нетональных компонентов сигнала.

MP3 остается популярным и широко используемым форматом благодаря своей совместимости с большим количеством устройств и программного обеспечения, но AAC предлагает лучшие возможности и становится предпочтительным выбором для новых приложений, требующих высокой эффективности сжатия и качества звука.

Вывод

Сравнение различных характеристик аудиокодеков

Характеристика	MP3	AAC
Год выпуска	1993	1997
Эффективность сжатия	Средняя	Высокая

Характеристика	MP3	AAC
Поддержка битрейтов	До 320 kbps	До 320 kbps и выше
Скорость кодирования	Высокая	Высокая
Качество при низком битрейте	Удовлетворительное	Высокое
Совместимость и распространённость	Очень высокая	Высокая
Лицензирование	Требует уплаты роялти	Требует уплаты роялти
Энтропийное кодирование	Huffman coding	Huffman coding
Адаптивная фильтрация	Нет	Использование более сложных фильтров

Где какой аудиокодек использовать и почему

- **MP3**
 - **Применение:** Музыкальные файлы, подкасты, аудиокниги, передача аудио по интернету.
 - **Почему:** Высокая совместимость с различными устройствами и платформами, быстрая скорость кодирования, простота использования и повсеместное распространение.
- **AAC**
 - **Применение:** Стриминговые сервисы (Apple Music, YouTube), телевидение, высококачественные аудиофайлы.
 - **Почему:** Высокая эффективность сжатия, лучшее качество при низких битрейтах, поддержка современных устройств и приложений, превосходит MP3 по качеству звука при тех же битрейтах.

Заключение

Выбор аудиокодека зависит от конкретных требований и условий использования.

- **MP3** остается наиболее распространенным аудиокодеком благодаря своей высокой совместимости и быстрой скорости кодирования. Он подходит для широкого спектра приложений, от музыкальных файлов до подкастов.
- **AAC** обеспечивает лучшую эффективность сжатия и качество звука, что делает его предпочтительным для современных приложений, таких как стриминговые сервисы и телевидение. Однако его использование может быть ограничено из-за лицензионных сборов.

Хотя MP3 и AAC имеют свои сильные стороны, AAC становится все более популярным благодаря своему лучшему качеству звука и поддержке современных технологий.