**Psychoakustische Grundlagen**

**Hörschwelle**

Die Hörschwelle ist die minimale Lautstärke, bei der ein bestimmter Ton noch wahrgenommen wird. Dies kann unterschiedlich sein, je nach Frequenz des Tons.

**Fletcher-Munson equal-loudness curves**

Dies sind Kurven, die die Beziehung zwischen der Lautstärke und der Frequenz von Tönen darstellen, bei denen sie als gleich laut empfunden werden. Sie zeigen, dass niedrigere Frequenzen bei niedrigeren Lautstärken wahrgenommen werden müssen als höhere Frequenzen.

**Lautstärkepegel / Phon**

Der Lautstärkepegel ist ein Maß für die Lautstärke eines Tons. Es wird in Dezibel (dB) gemessen. Phon ist eine spezielle Einheit des Lautstärkepegels, die an die Fletcher-Munson equal-loudness curves angepasst ist.

**Frequency Masking**

Dies beschreibt das Phänomen, dass ein Ton in einer bestimmten Frequenz die Wahrnehmung von Tönen in benachbarten Frequenzen beeinträchtigen kann. Niedrigere Töne können höhere Töne gänzlich unhörbar machen, während es umgekehrt nicht ganz so effektiv ist.

**Just Noticeable Difference / Frequenzgruppen**

Just Noticeable Difference (JND) beschreibt die minimale Veränderung einer bestimmten Eigenschaft (z.B. Lautstärke), die noch wahrgenommen wird. Frequenzgruppen beziehen sich darauf, dass bestimmte Frequenzbereiche (z.B. Bass, Mitte, Höhen) unterschiedlich empfindlich auf Veränderungen reagieren.

**Temporal Masking (pre-/post-masking)**

Temporal Masking beschreibt das Phänomen, dass ein Ton die Wahrnehmung von Tönen beeinflussen kann, die vor oder nach ihm auftreten. Pre-masking bezieht sich auf die Beeinflussung von Tönen, die vor dem aktuellen Ton auftreten, post-masking auf die Beeinflussung von Tönen, die nach dem aktuellen Ton auftreten.

**Welcher Ton wird von einem Menschen als lauter wahrgenommen:**

1. Ein Ton mit 100Hz und 80dB ~62Phon
2. Ein Ton mit 1.000Hz und 80dB ~80Phon
3. Ein Ton mit 10.000Hz und 80dB ~68Phon

**Wie laut nimmt ein Mensch einen Ton mit 4.000Hz und 80dB wahr?**

Ca 80 Phon

**Wie laut nimmt ein Mensch einen Ton mit 60Hz und 60dB wahr?**

Ca 20 Phon

**MP3 / MPEG-Audio**

**In welchem Zusammenhang steht MP3 mit MPEG-1?**

MP3 ist Teil des MPEG-1 (Moving Picture Experts Group) Standards. MPEG-1 ist ein internationaler Standard für die Kompression von digitalen Audiosignalen und Video-Signalen und wurde von der ISO (International Organization for Standardization) und IEC (International Electrotechnical Commission) entwickelt.

MPEG-1 definiert drei Schichten für die Kompression von Audio- und Video-Signalen:

* Layer I: Eine niedrigere Qualität, aber höhere Kompression
* Layer II: Eine mittlere Qualität und Kompression
* Layer III: Eine höhere Qualität, aber geringere Kompression

MP3 ist die Bezeichnung für die dritte Schicht (Layer III) des MPEG-1 Standards und wurde speziell für die Kompression von Audiosignalen entwickelt. MP3 nutzt die perceptuellen Eigenschaften des menschlichen Hörvermögens, um Audiosignale zu komprimieren, indem es die Frequenzen entfernt, die das menschliche Ohr nicht so empfindlich wahrnimmt.

**Wie kann MP3 die sehr hohe Kompressionsrate trotz guter bzw. akzeptabler Qualität erreichen?**

**Erklären Sie den prinzipiellen Aufbau eines MP3-Encoders anhand einer Grafik.**

**Welche Informationen werden vom Encoder verworfen und warum haben diese relativ wenig Einfluss auf die wahrgenommene Qualität der resultierenden Audiodatei.**

**Handelt es sich bei MP3 um einen Perceptual audio codec? Was versteht man darunter?**

Ja, MP3 ist ein perceptueller Audio-Codec.

Ein perceptueller Codec nutzt die Tatsache, dass das menschliche Hörvermögen bestimmte Frequenzen und Lautstärken nicht so empfindlich wahrnimmt wie andere. Es komprimiert die Audiodaten, indem es die Informationen entfernt, die das menschliche Ohr nicht wahrnehmen kann, ohne die Qualität des Audios wesentlich zu beeinträchtigen.

**Recherchieren Sie die Eigenschaften von WAV Dateien.**

**Unkomprimiert**

WAV-Dateien sind unkomprimiert, was bedeutet, dass sie die Originaldaten des Audios ohne jegliche Datenreduktion enthalten. Dies führt zu einer hohen Audioqualität, aber auch zu großen Dateigrößen.

**Breite Unterstützung von Audio-Formats**

WAV-Dateien unterstützen eine breite Palette an Audio-Formats einschließlich Monofonie und Stereophonie, sowie eine große Anzahl an Bitraten und Sampling-Frequenzen.

**Metadaten**

WAV-Dateien haben die Möglichkeit, Metadaten wie Künstler, Album und Titel zu enthalten.

**Kompatibilität**

WAV-Dateien sind weit verbreitet und kompatibel mit den meisten Betriebssystemen und Audio-Software.

**Nutzen Sie FFmpeg um die Datei in das MP3 Format zu konvertieren. Erzeugen Sie MP3 Dateien in den Bitraten:**

32 kbps ffmpeg -i in.wav -b:a 32k output.mp3

64 kbps ffmpeg -i in.wav -b:a 64k output.mp3

128 kbps ffmpeg -i in.wav -b:a 128k output.mp3

320 kbps. ffmpeg -i in.wav -b:a 320k output.mp3

Können Sie Unterschiede feststellen?

**Advanced Audio Coding (AAC)**

**Wo wird AAC hauptsächlich eingesetzt?**

MP3 ist sehr weit verbreitet in Musik-Player, Auto-Stereoanlagen und in PCs, während AAC ist sehr weit verbreitet in Streaming-Dienste wie Spotify, Apple Music und YouTube.

**Beschreiben Sie wie AAC grundsätzlich funktioniert und skizzieren Sie die einzelnen Schritte des Encoders.**

**Listen Sie die wichtigsten Unterschiede zwischen MP3 und AAC auf.**

**Audioqualität**

AAC bietet in der Regel eine bessere Audioqualität als MP3 bei gleicher Bitrate. AAC nutzt fortgeschrittene Techniken wie die psychoacoustic Modelle, die es ermöglichen, Audiosignale noch effektiver zu komprimieren.

**Frequenzbereich**

AAC unterstützt einen breiteren Frequenzbereich als MP3, was bedeutet, dass es höhere und tiefere Frequenzen besser darstellen kann.

**Kompatibilität**

MP3 ist weit verbreitet und kompatibel mit den meisten Betriebssystemen und Audio-Software, AAC ist jedoch auch weit verbreitet, aber nicht so weit verbreitet wie MP3.

**Erzeugen Sie MP3 Dateien in den Bitraten:**

32 kbps ffmpeg -i in.wav -c:a libfdk\_aac -b:a 32k output.aac

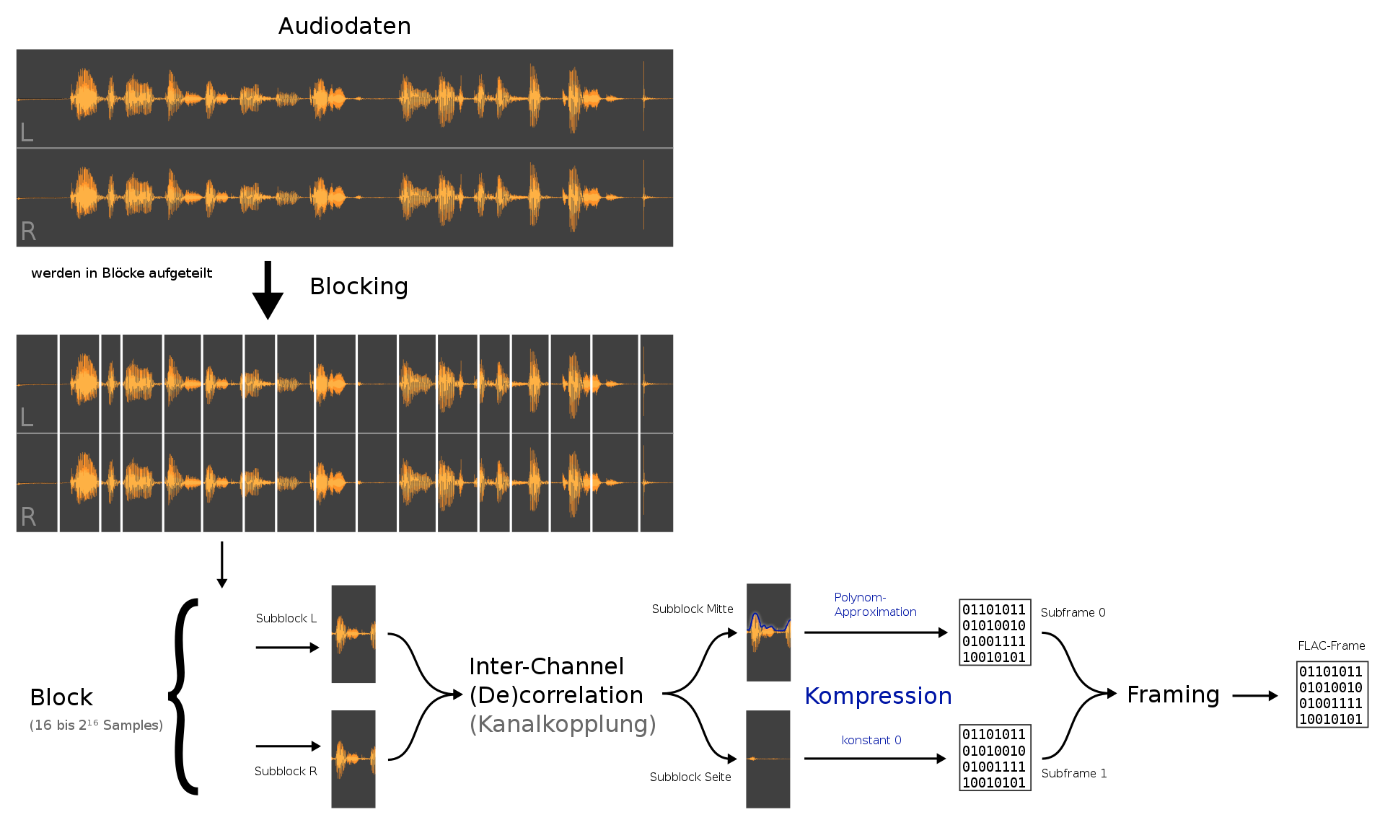
64 kbps ffmpeg -i in.wav -c:a libfdk\_aac -b:a 32k output.aac

128 kbps ffmpeg -i in.wav -c:a libfdk\_aac -b:a 32k output.aac

320 kbps. ffmpeg -i in.wav -c:a libfdk\_aac -b:a 32k output.aac

**Free Lossless Audio Codec (FLAC)**

**Recherchieren Sie die Funktionsweise des Codecs im Internet.**



**Blocking**

FLAC unterteilt die Audiodaten stets in Blöcke zu je 16 bis 65535 () Samples. Die Subblöcke, also die verschiedenen Kanäle eines Blocks, werden gemeinsam in einem Frame untergebracht, d. h. komprimiert und enkodiert.

**Inter-Channel Decorrelation**

Sofern ein Stereo-Signal anliegt, kann es aus der vorliegenden Links-Rechts-Kodierung (d. h. jeder Kanal ist separat codiert) in eine Mid-Side-Kodierung überführt werden. Dies geschieht entweder fest (also immer bei L/R belassen bzw. immer in M/S umwandeln) oder adaptiv in jedem Frame.

**Modeling**

Der Werteverlauf jedes Subblocks wird entweder mittels einer Polynomfunktion oder mit dem Verfahren Linear Predictive Coding angenähert. Einige übliche Kombinationen an Koeffizienten sind verfügbar, oder diese werden direkt im Frame gespeichert.

**Residual Coding**

Das Fehlersignal, also der Unterschied zwischen dem tatsächlichen Signal und dem modellierten Signal, wird mittels Rice-Kodierung verlustfrei im Subframe gespeichert.

**Framing**

Zuletzt werden die entstandenen Frames mit Header und Footer versehen, die unter anderem für Fehlererkennung mittels CRC und Synchronisierbarkeit sorgen.

**Finden Sie heraus, wie gut die durchschnittliche Kompressionsrate von FLAC ist und wovon die tatsächlich erreichte Kompressionsrate abhängt.**

Laut Angaben der Entwickler erreicht FLAC durchschnittlich eine Komprimierung auf rund 50 Prozent der Ausgangsgröße. Es kommt auf die Blöckgröße und die LPC-Koeffizienten an.

**Nutzen Sie Ffmpeg oder den FLAC Encoder (https://xiph.org/flac/) um aus der WAV-Datei in.wav FLAC-Dateien zu erzeugen.**

ffmpeg -i input.wav -c:a flac output.flac

**FLAC bietet 9 verschiedene Kompressionsstufen. Finden Sie heraus, worin sich diese Kompressionsstufen unterscheiden.**

Die Stufen 0 bis 2 arbeiten mit recht kleinen Blöcken und fixen LPC-Koeffizienten, um die Kodiergeschwindigkeit zu erhöhen.

Die Stufen 3 bis 6 benutzen größere Blöcke und einen LPC-Grad bis zu acht, um platzsparender zu komprimieren.

Die Stufen 7 und 8 versuchen zusätzlich, durch ausgiebige Suche nach den optimalen Kodierungsparametern (exhaustive model search) die Zieldatei auf eine minimale Größe zu komprimieren, was die Enkodiergeschwindigkeit deutlich verlangsamt.