

Audioformate

In diesem Kapitel werden Möglichkeiten zum Codieren von Audiodateien beschrieben. Auch hier gilt, dass die beschriebene Auswahl nur die am weitest verbreiteten Verfahren umfasst. Generell muss man zwischen verlustfreier und verlustbehafteter Kompression unterscheiden.

Verlustfreie Kompression

Verglichen mit Bildkomprimierung werden verlustfreie Komprimierungsmethoden in der Audiokomprimierung nicht so häufig verwendet. Die Hauptbenutzer von verlustfreier Komprimierung sind Toningenieure und ihre Kunden, die den Qualitätsverlust der verlustbehafteten Kompressionsmethoden vermeiden wollen. Der Grund für die geringe Verwendung ist die geringe Speicherplatzreduktion, die hierbei erreicht werden kann.

Die meisten Tonaufnahmen sind Sounds aus der realen Welt. Solche Daten können nicht gut komprimiert werden. Aber auch computergenerierte Tonabfolgen können sehr komplizierte Wellenformen enthalten, die sich mit vielen Kompressionsalgorithmen nur schlecht verkleinern lassen. Dies liegt an der Natur der Schallwellen, die sich im Allgemeinen schwer vereinfachen lassen. Vereinfachungen führen dazu, dass nicht mehr alle Frequenzen des originalen Signals repräsentiert werden und so bereits ein Qualitätsverlust auftritt. Außerdem ändern sich die Werte der Audiosamples sehr schnell und es gibt selten Folgen von gleichen Bytes, weswegen allgemeine Datenkompressionsalgorithmen (wie Lauflängencodierung) keine hohe Datenreduktion erreichen.

Die verlustfreien Audiocodecs unterscheiden sich von allgemeinen Algorithmen zur Datenreduktion dadurch, dass sie speziell an die typische Datenstruktur von Audiodateien angepasst sind. So können Audiocodecs Ähnlichkeiten zwischen den Kanälen (links, rechts) oder eine bestimmte Samplerate (z.B. CD-Qualität: 44,1 kHz) als Voraussetzung für die zu komprimierenden Daten annehmen. Dies hat zur Folge, dass die Audiocodecs in fast allen Fällen Audiodateien besser komprimieren als z. B. die ZIP- oder RAR-Algorithmen. Die Kompressionsrate von verlustfreien Audiocodecs bei CD-Tracks liegt meist zwischen 40 und 70 Prozent, aber meistens näher bei 40 Prozent.

Verlustfreien Codecs reproduzieren jeweils das Originalsignal.

Apple Lossless

Apple Lossless ist ein von Apple entwickelter proprietärer Codec zur verlustfreien Audio-kompression. Ab Version 4.5 von iTunes¹, kann er benutzt werden, um Musik von CDs und anderen Quellen (AIFF, WAV) zu importieren, wobei die komprimierten Audiodaten in MP4-Container (MP4 ist der für MPEG-4-Inhalte von der Moving Picture Experts Group vorgesehene Container.) eingebettet werden. Apple Lossless wird auch von Apples AirPort Express verwendet, um Musik mit iTunes 4.6 und AirTunes² über WLAN zu streamen.

¹ iTunes ist ein kostenloser Musikplayer, sowohl für den Computer der Firma Apple als auch jeden anderen Computer, der unter dem System Windows 2000 bzw. XP läuft. Die Funktionen erstrecken sich vom Abspielen über das Organisieren bis hin zum Kaufen von Musik. Die aktuelle Versionsnummer ist 4.7.1, die für die Betriebssysteme Mac OS X sowie für Windows 2000 und XP erhältlich ist. Die Mac-OS-9-Version wird seit der Version 2.0.4 nicht mehr weiterentwickelt.

² AirTunes nennt sich die Methode, über Wireless-Lan Musik mit der Software iTunes über das Gerät Airport Express der Firma Apple zu streamen. Dabei wird in der Software iTunes die Airport Express Station als Lautsprecher ausgewählt und anschließend Musik abgespielt. An die Airport Express Station wird über einen Lineout Ausgang direkt die Stereoanlage oder Ähnliches angeschlossen.

Free Lossless Audio Codec

FLAC, ausgeschrieben Free Lossless Audio Codec, zu Deutsch Freier verlustfreier Audio-Kodierer. Frei bedeutet, dass es sich um Freie Software handelt. Ähnlich wie MP3 komprimiert FLAC Musikdaten, aber mit dem Unterschied, dass dies ohne Qualitätsverluste geschieht.

Verlustbehaftete Kompression

Die meisten verlustbehafteten Kompressionsalgorithmen basieren auf simplen Transformationen, wie der „modifizierten diskreten Kosinus-Transformation“ (MDCT), welche die aufgenommene Wellenform in ihre Frequenzabfolgen umwandeln. Einige moderne Algorithmen benutzen Wavelets, aber es ist noch nicht sicher, ob solche Algorithmen besser funktionieren (bessere Qualität/Kompressionsverhältnisse erzielen) als die auf MDCT basierenden.

Die meisten Algorithmen versuchen nicht den mathematischen Fehler zu reduzieren, der beim Abtasten eines Signals entsteht, sondern die subjektive menschliche Wahrnehmung der Tonfolgen zu verbessern. Da das menschliche Ohr nicht alle Informationen, die in einem Sound enthalten sind, analysieren kann, ist es möglich, eine Sounddatei stark zu verändern, ohne dass die subjektive Wahrnehmung des Hörers beeinträchtigt wird. So kann ein Codec zum Beispiel einen Teil der sehr hohen und sehr tiefen Frequenzen (welche für Menschen fast unhörbar sind) weglassen. Auf ähnliche Weise werden Frequenzen, die durch andere Frequenzen überlagert sind, mit geringerer Genauigkeit wiedergegeben. Eine andere Art der Überlagerung ist, dass ein leiser Ton nicht erkennbar ist, wenn er unmittelbar vor oder nach einem lauten Ton kommt. Ein solches Modell der Ohr-Gehirn Verbindung, welches für diese Effekte verantwortlich ist, wird häufig psychoakustisches Modell genannt (auch: „Psychoacoustic Model“, „Psycho-model“ oder „Psy-model“). Ausgenutzt werden hierbei Eigenschaften des menschlichen Gehörs wie Frequenzgruppenbildung, Hörbereichsgrenzen, Maskierungseffekte, sowie das Wissen über die Signalverarbeitung des Innenohrs.

Aufgrund der Natur der verlustbehafteten Algorithmen verschlechtert sich die Qualität, wenn eine solche Datei dekomprimiert und anschließend wieder komprimiert wird (Generationsverluste). Das passiert in der Praxis vor allem, wenn eine Audio-CD aus verlustbehafteten Audiodateien gebrannt wird und diese später wieder komprimiert wird. Dies macht verlustbehaftete Dateien sehr ungeeignet für Anwendungen in professionellen Tonbearbeitungsbereichen. Allerdings sind solche Dateien sehr beliebt bei Endbenutzern, da ein Megabyte je nach Komplexität des Tonmaterials ungefähr für eine Minute Musik bei annehmbarer Qualität reicht, was einer Kompressionsrate von ca. 1:11 entspricht.

MP2

MPEG-1 Audio Layer 2 (MP2) ist ein auch als Musicam bekanntes und von der Motion Pictures Experts Group standardisiertes Verfahren zur Audiokompression von digitalen Audiodaten. Die Abkürzung wird auch als Dateierweiterung für Dateien, die Audiodaten dieses Typs beinhalten, genutzt.

Während es im PC- und Internet-Bereich im Wesentlichen durch MP3 abgelöst wurde, verbleibt es als dominanter Standard für digitalen Rundfunk als Teil der DAB- und DVB-Standards.

MP3

MP3, eigentlich MPEG-1 Audio Layer 3, ist ein Dateiformat zur verlustbehafteten Audiokompression. Entwickelt wurde es 1987 am Fraunhofer-Institut für Integrierte Schaltungen in Erlangen in Zusammenarbeit mit AT&T Bell Labs und Thomson. Wie viele der aktuellen Kodierverfahren sind Kernbereiche von MP3 durch Patente geschützt. Prof. Dr. Karlheinz

Brandenburg wurde für die Entwicklung dieses Datenformates mehrfach ausgezeichnet. Brandenburg ist heute Leiter des Fraunhofer-Instituts für Digitale Medientechnologie IDMT in Ilmenau und Direktor am Institut für Medientechnik der TU Ilmenau.

Wie andere Formate setzt MP3 darauf, dass die Wahrnehmung des Menschen begrenzt ist. Die Menge der Töne, die vom Menschen beispielsweise wegen ihrer Frequenz oder Lautstärke nicht wahrgenommen werden können, wird beim MP3 Verfahren reduziert. Somit ist MP3 ein verlustbehaftetes Verfahren, d.h. das Eingangssignal kann aus dem Ausgangssignal nicht mehr reproduziert werden und Informationen gehen verloren. Die Verluste hängen stark von der Datenrate und dem zu komprimierenden Tonmaterial ab. Das MP3-Format erlaubt dem Benutzer, Kompressionsraten von 32 kBit/s bis zu 320 kBit/s festzulegen. Üblich ist für Musik, die in lauter Umgebung oder bei hoher Lautstärke gehört wird, eine Bitrate von 128 kBit/s, je nach Anspruch auch eine höhere. Weiterhin gibt es variable Bitraten, die sich einer gewählten Qualitätsstufe anpassen.

MP3 wurde vor allem durch Musiktäuschbörsen in der breiten Öffentlichkeit bekannt, wird aber auch bei vielen DVD-Rips als Audioformat benutzt. Die MP3-Technologie wird für so genannte MP3-Player eingesetzt, mit denen man auch unterwegs Musik hören kann.

Advanced Audio Coding

Advanced Audio Coding, häufig auch einfach AAC genannt, ist ein von der Moving Picture Experts Group (Dolby, Fraunhofer-Institut für Integrierte Schaltungen in Erlangen, AT&T, Nokia, Sony) entwickeltes Audiokompressionsverfahren, das vom MPEG-2 Standard spezifiziert ist. Es bietet gegenüber älteren Formaten wie MP3 diverse Vorteile, unter anderem eine bessere Qualität bei gleicher Dateigröße oder die Unterstützung für Multichannel-Audio. Jedoch benötigen manche Profile (Kompressionseinstellungen) mehr Rechenleistung zum Encodieren (zum Beispiel beim Rippen von CDs) und Dekodieren (Abspielen) als zum Beispiel MP3.

Neben dem Kompressionsschema ist auch das Containerformat MP4 definiert worden, das unter anderem die Übermittlung von Metadaten (Tagging) oder die Verwendung eines Kopierschutzverfahrens (siehe DRM) erlaubt.

Windows Media Audio

Windows Media Audio (WMA) ist ein proprietärer Audio-Codec von Microsoft und Teil der Windows Media-Plattform. WMA ist ein verlustbehafteter Codec, d.h. beim Kodierungsprozess gehen Daten verloren. In der „Professional“-Variante unterstützt er bis zu 24 bit/96 KHz bei einer variablen Bitrate von bis zu 768 kb/s und Surround-Ton mit bis zu 7.1 Kanälen. Daneben gibt es eine Version, die speziell auf Quellmaterial, das Stimmaufnahmen enthält, abgestimmt ist (Windows Media Audio Voice, ACELP), sowie den verlustfreien Codec Windows Media Audio Lossless.

WMA-kodierte Audio-Streams, sind in der Regel in ASF-Container eingebettet. Im Falle einer reinen Audio-Datei tragen die Dateinamen zumeist die Endung .wma. Windows Media Audio unterstützt auch die Einbindung von DRM, die dem Urheber von geistigem Eigentum die Regelung der Zugriffsrechte auf Tonmaterial ermöglichen soll. In der Praxis wird diese Technik häufig als Kopierschutzmaßnahme bzw. als Abspielbeschränkung eingesetzt.

Adaptive Transform Acoustic Coding

ATRAC ist die Abkürzung für Adaptive Transform Acoustic Coding (in etwa: Anpassungsfähiger Ton-Umwandlungscode). ATRAC ist ein Audiokompressionsverfahren, das auf psychoakustischen Grundregeln basiert. Es komprimiert Audio-Dateien auf einen Bruchteil des

ursprünglichen Datenvolumens. Der ATRAC-Codec wurde von Sony mit der Einführung der MiniDisc im Jahre 1992 vorgestellt und bis heute stetig weiterentwickelt.

Der Codec beinhaltet verschiedene Kopierschutzmechanismen. Seit ATRAC3 wird auch Digital Rights Management unterstützt. Sony möchte so Raubkopien unterbinden, schränkt aber damit auch die Nutzungsmöglichkeiten des ATRAC-Codex ein und hat so dessen großen Durchbruch bisher verhindert. Vor allem wegen der hohen Lizenzgebühren, die von Sony für ATRAC verlangt werden, findet man heute nur noch wenige Firmen, die kompatible Geräte herstellen. Darunter sind Sony selbst, Sharp, Aiwa (von Sony aufgekauft) und Panasonic. Trotz alledem erfreuen sich ATRAC-kompatible Geräte und die unweigerlich mit diesem Codec verknüpfte MiniDisc großer Beliebtheit.

Digital Theater Systems

DTS ist ein Mehrkanal-Tonformat, das im Kino, bei Laserdisks und DVDs zum Einsatz kommt. DTS verwendet verlustbehaftete Verfahren zur Komprimierung der Audiodaten. Es wurde von der Firma Digital Theatre Systems Inc. entwickelt. DTS ist die Abkürzung für Digital Theatre Sound und wird als höherwertige Alternative zu Dolby Digital vermarktet.

Die Bitrate von DTS Tonspuren auf Video DVDs beträgt entweder 754.5 kbps oder 1509.25 kbps. Die Tonspur kann einen LFE-Kanal enthalten, der nur den für die Tieftonwiedergabe zuständigen Subwoofer bedient. Die vollfrequenten Kanäle gehen von 20Hz bis 20kHz, der Basskanal reicht aber nur bis 80Hz.

Der Ton wird im Kino auf einer oder mehreren Compact Discs gespeichert, die durch eine kleine Timecode-Spur auf der Filmrolle zum Bild synchronisiert werden. Der analoge Ton befindet sich weiterhin auf dem Film und wird bei Ausfall von DTS als Notlösung verwendet.

Dolby Digital

Dolby Digital (auch DD, SR*D oder AC-3 genannt) ist ein Mehrkanal-Tonsystem, das im Kino, auf Laserdisks, DVDs und im Fernsehen zum Einsatz kommt. Es konkurriert im Bereich Kino und DVD mit den Verfahren DTS und SDDS (SDDS nur Kino).

Dolby Digital ermöglicht die datenreduzierte Kodierung von drei Frontkanälen, zwei vollwertigen Rückkanälen und einem LFE-Kanal (Low Frequency Effect; Subwoofer) in einen konstanten Datenstrom. Das verwendete Kodierungsverfahren heißt AC-3. Die vollfrequenten Kanäle haben ein Spektrum von 20Hz bis 20kHz, der Basskanal reicht bis 120Hz.

Im Heimbereich ist Dolby Digital der digitale Nachfolger von Dolby Surround (analoges Mehrkanal-Tonsystem). Zusätzlich enthält es Metadaten, die dem Dekoder u.a. behilflich sind, einen Stereo-Downmix der sechs verwendeten Kanäle zu erstellen und den Dynamikbereich (Differenz zwischen maximaler und minimaler Amplitude), falls dies der Zuhörer möchte, zu beschränken.