SLUB
Wir führen Wissen.

Handreichung retrodigitalisierte Tondokumente

SLUB Dresden

Version 1.0, 2023-09-22

Inhaltsverzeichnis

Vorwort . . . 1
Allgemeines . . . 1
Abgrenzung . . . 2
Theoretische Grundlagen . . . 3
- Abtastrate . . . 3
- Quantisierungsrauschen . . . 3
- Wiedergabetreue . . . 4

Anforderungen . . . 5
- Allgemeine Anforderungen . . . 5
- begleitende Metadaten . . . 7

Validierung . . . 7
Glossar . . . 8
Bibliographie . . . 9

Release Note 1.0 vom 2023-09-22

- Initales Dokument

Vorwort

Dieses Dokument richtet sich an Produzenten, die digitale Objekte in das SLUBArchiv.digital einliefern und diese langfristig benutzbar erhalten wollen.

Allgemeines

Analoge Tondokumente zu sichern, ist eine Herausforderung für Gedächtnisorganisationen.

Eine Möglichkeit ist die Retrodigitalisierung und anschliessende digitale Langzeitarchivierung von vom Verfall bedrohten Tondokumenten. Die Hauptvorteile liegen in der einfachen Weiterverwendung und der möglichen, verlustfreien Kopierbarkeit der Digitalisate.

In den letzten Jahren hat sich herauskristallisiert, dass Free Lossless Audio Codec (FLAC) [https://xiph.org/flac] (als verlustfreier Audiocodec) Eigenschaften mitbringt, die die Langzeitarchivierungs-Community zu schätzen weiß:

- offene Implementierung durch Referenzimplementierung [https://xiph.org/flac/download.html] und die weit verbreiteten FFMPEG- und libAV-Softwarebibliotheken
- offene Standardisierung im Rahmen der Cellar working group der Internet Engineering Task Force (IETF)
- hohe Kodiereffizienz
- gute Robustheit durch integrierte CRC-Summen
- verfügbare Werkzeuge, insbesondere quelloffener Validator MediaConch [https://mediaarea.net/MediaConch/] im Rahmen des EU-Projektes Preforma

Die SLUB hat sich daher für die Ablieferung von Tondokumenten an das SLUBArchiv auf Free Lossless Audio Codec (FLAC) [https://xiph.org/flac] als Dateiformat [1] festgelegt.

Die erlaubten Parameter, die für die Ablieferung verwendet werden dürfen, werden im Folgenden beschrieben. Als Basis dient eine interne Analyse der Nutzungsszenarien und der daraus abgeleiteten signifikanten Eigenschaften für den Objekttyp Tonband.

Vorbehaltlich gesonderter Absprachen mit dem SLUBArchiv sind von dieser Handreichung abweichende Änderungen **nicht** gestattet. Informationspakete, die diese Anforderung nicht erfüllen, sind nicht langzeitarchivfähig und werden vom SLUBArchiv zurückgewiesen.

Abgrenzung

Dieses Dokument beschreibt die Anforderungen des SLUBArchiv an retrodigitalisierte Tondokumente, die auf Magnetband (Tonbandrollen, Audiokasette) vorliegen.

Nicht von diesem Dokument erfasst [2] sind Audiomaterialien,

- die aus digitalen Vor- und Nachlässen stammen
- die originär digital erzeugt wurden
- die aus Videostreams aufgezeichnet wurden
- die Bestandteile von retrodigitalisiertem Film - und Video darstellen
- die Teile von Kunstinstallationen oder musealen Charakters sind
- die aus Produktionsarchiven von Sendeanstalten stammen
- die Hochgeschwindigkeits- oder Zeitlupenaufnahmen aus dem wissenschaftlichen Bereich enthalten

Dieses Dokument ist nicht auf Präsentations- oder Masternutzungsdateien (Mezzanine-Dateien) anzuwenden, da diese **grundsätzlich** nicht in das SLUBArchiv aufgenommen werden.

Digitalisierungsanforderungen an Dienstleister sind im Einklang mit diesem Dokument zu verfassen.

Theoretische Grundlagen

Die Herausgabe von Empfehlungen, insbesondere im A/V Bereich, sorgt immer wieder für Diskussionen. Um diese zu versachlichen, werden die dieser Handreichung zugrundeliegenden Entscheidungen technisch-physikalisch begründet und die wissenschaftlichen Quellen angegeben.

Abtastrate

Die Grundlage der Digitalisierung ist mit Shannon (1984) gelegt:

> "A method is developed for representing any communication system geometrically. Messages and the corresponding signals are points in two "function spaces," and the modulation process is a mapping of one space into the other. Using this representation, a number of results in communication theory are deduced concerning expansion and compression of bandwidth and the threshold effect. Formulas are found for the maximum rate of transmission of binary digits over a system when the signal is perturbed by various types of noise. Some of the properties of "ideal" systems which transmit at this maxmum rate are discussed. The equivalent number of binary digits per second for certain information sources is calculated."

Laut Shannon gilt:

$$f_a > 2 f_{max}$$

Das heißt, dass ein mit mit der Abtastfrequenz f_a digitalisiertes analoges Signal komplett rekonstruierbar ist, wenn die Abtastfrequenz mindestens doppelt so groß ist wie die höchste im analogen Signal vorkommende Frequenz f_{max}.

Die Ergebnisse werden von Unser (2000) auch unter Verwendung nicht idealer Antialiasing-Filter bestätigt:

> "[...]We then extend the standard sampling paradigm for a presentation of functions in the more general class of "shift-in-variant" function spaces, including splines and wavelets. Practically, this allows for simpler-and possibly more realistic-interpolation models, which can be used in conjunction with a much wider class of (anti-aliasing) prefilters that are not necessarily ideal low-pass. We summarize and discuss the results available for the determination of the approximation error and of the sampling rate when the input of the system is essentially arbitrary; e.g., nonbandlimited.[...]"

Da alle Tonbandaufzeichnungen im Kontext dieser Handreichung üblicherweise nur Signale im hörbaren Bereich enthalten und weiterhin anzunehmen ist, dass die Verarbeitungskette bei der Aufnahme (Tonaufnehmer und Mikrophone, Verstärker) und bei der Digitalisierung für den Hörbereich eingerichtet sind, können Frequenzen über 20kHz im Regelfall ausgeschlossen werden.

Eine Übersicht zu den Grenzfrequenzen typischer Medien findet sich in Webers (1974).

Quantisierungsrauschen

Die Formel für Quantisierungsrauschen lautet nach Bennett (1948):

$$Q_{SNR} = (6,\ 02n + 1,\ 76)\text{dB}$$

Sprich eine Quantisierungsauflösung von 16Bit ergibt demnach einen Signal-Rauschabstand von 98,1dB. Dieser Wert liegt deutlich über dem Signal-Rauschabstand der typischerweise verwendeten Geräte (Verstärker, Tonbandgeräte), so dass das Quantisierungsrauschen im Rauschen der Geräte unter geht.

Dadurch, dass in den meisten Fällen überabgetastet wird, erhöht sich der Signal-Rausch-Abstand (vgl. Kester (2022), Stichwort: process gain) um den additiven Anteil von

$$dQ_{SNR} = 10\log_{10}\frac{f_a}{2f_{max}}$$

Demnach ist eine generelle Erhöhung der Quantisierungauflösung auf 24Bit zum aktuellen Stand nicht sinnvoll, kann in begründeten Fällen aber genutzt werden.

Wiedergabetreue

In Meyer & Moran (2007) wurde untersucht, ob Testpersonen Audiomaterial, welches in 16-bit, 44,1kHz digitalisiert wurde, von analogem unterscheiden können:

> "The authors report on a series of double-blind tests comparing the analog output of high-resolution players playing high-resolution recordings with the same signal passed through a 16-bit/44.1-kHz “bottleneck.” [...] The test results show that the CD-quality A/D/A loop was undetectable at normal-to-loud listening levels, by any of the subjects, on any of the playback systems."

Anforderungen

Allgemeine Anforderungen

Verpflichtend

- Alle Digitalisate müssen als Free Lossless Audio Codec (FLAC) [https://xiph.org/flac] Dateien abgeliefert werden, sh. Free Lossless Audio Codec Format [https://datatracker.ietf.org/doc/draft-ietf-cellar-flac/]
- Als Datei- und Verzeichnisnamen sollte man sich auf den Vorrat A-Z, a-z, 0-9 und die Sonderzeichen "-._" beschränken. Dies erleichtert den Kopiervorgang über Betriebssystemgrenzen hinweg.
- Der grundsätzliche Aufbau von SIPs zur Einlieferung der digitalen Objekte ist in der SIP-Spezifikation beschrieben.

Gute Praxis

- Für die materialspezifisch korrekte Wahl der qualitätsbestimmenden Digitalisierungsparameter Abtastfrequenz und Quantisierungsauflösung sollten die Empfehlungen von IASA, memoriav oder vergleichbarer Institutionen Berücksichtigung finden.
- Es hat sich bewährt, je eine Datei pro analogem Originaldatenträger (Tonband) und Durchlauf anzulegen. Das heißt:
  - Bei Datenträgern, die doppelseitig bespielt sind, wird je eine Datei pro Seite angelegt
  - Stereoaufnahmen, die auf zwei Kanälen in einem Durchlauf aufgenommen wurden, werden zusammen in einer Datei kodiert.
- Für sonstige Aufnahmen im Mehrspurverfahren gilt, dass je eine Datei pro Spur angelegt wird.
- Im Rückwärtsbetrieb digitalisierte Spuren sind spielrichtig abzuspeichern.
- Kapitelmarken sollten verwendet werden, um auf den Beginn verschiedener Inhalte der Originalmedien zu verweisen. Dies ist dann notwendig, wenn zB. auf einem Tonband unterschiedliche Werke aufgezeichnet wurden.
- Auf eine Verwendung von verlustbehafteten Zwischenformaten (Audiocodecs) bei der Digitalisierung und anschliessenden Wandlung nach FLAC ist zu verzichten. Wenn möglich, sollte nativ nach FLAC digitalisiert werden. Unter https://xiph.org/flac/links.html#hardware findet sich eine Übersicht unterstützter Hardware.

Verpflichtend

Codec

Durch die Verwendung von FLAC als Dateiformat wird auf den lossless Codec FLAC zurückgegriffen.

Kodierungsparameter

Für Audiodigitalisierung [3] werden folgende Parameter verwendet:

- linear PCM (little endian)
- Quantisierungsauflösung: 16Bit pro Kanal (**ggf. 24Bit**, siehe unten)
- Samplingrate bzw. Abtastfrequenz (f_a): 48kHz (**ggf. Vielfache**, siehe unten)

Auch wenn die Materialien im Original Bandbreiten weit unterhalb der möglichen Grenzfrequenz von f_{max} = 24kHz aufweisen, wird aus folgenden Gründen 48kHz als Abtastfrequenz f_a gewählt:

- Normalisierung aller digitalisierten Tondokumente zur Vereinfachung der Processing-Workflows
- Sicherstellung, dass Filter von geringer Güte nicht zu Aliasingeffekten führen.

Im Zweifel sollte für die Abtastfrequenz sichergestellt sein, dass praktisch [4] gilt:

$$f_a > 2,1 f_{max}$$

In diesen Fällen sollte, um ein einfacheres Processing zu ermöglichen, als Abtastfrequenz ein ganzzahliges Vielfache von 48kHz verwendet werden.

Da eine Verdopplung der Abtastfrequenz zu einer Verdopplung der Datenmenge führt, sollte **nicht** pauschal mit mehr als 48kHz abgetastet werden!

Ähnliches gilt für eine Erhöhung der Quantisierungsstufen von 16Bit auf 24Bit, da oft nur das Rauschen, welches nicht informationstragend ist, in den unteren Quantisierungsstufen abgebildet wird.

Eine Abweichung von den vorgegebenen Werten ist durch den Produzenten materialspezifisch in den begleitenden Metadaten zu begründen.

Im Postprocessing können Audiodaten durch Upsampling mit ganzzahligem Faktor leicht auf ganzzahlige Vielfache von 48kHz angepasst werden. Dies ist zu berücksichtigen, damit bei der Digitalisierung höhere Abtastfrequenzen als für das Material zwingend notwendig vermieden werden.

begleitende Metadaten

Informationen, die für den Nutzer zum Verständnis der Qualität des Digitalisats beitragen, können als Sidecar-File als Text in UTF-8 (mit Zeilenumbruch *newline* ohne *carriage return*) gespeichert werden.

Sinnvoll sind Angaben zu:

- Begründung bei Abweichung von den Vorgaben dieses Dokuments
- Digitalisierungsvorbereitung (Reinigung, Reparatur)
- Digitalisierungsdurchführung (spezielle Abtastungsverfahren, Korrekturen)
- verwendete Digitizer und Parameter (Hersteller, Modell, Einstellungen)
- Software und Parameter (Hersteller, Versionsstand)
- Rechte (Bearbeitung, Schnitt…)
- Schnittdaten (Was wurde geschnitten und warum)
- Transkodierungsinformationen (verwendete Software und deren Parameter)
- Verwendetes Mehrspurverfahren

Validierung

Die Anforderungen des SLUBArchiv können mit der Software MediaConch [https://mediaarea.net/MediaConch/] geprüft werden und sind in den Profilen MediaConch-Profile [https://slubarchiv.slub-dresden.de/fileadmin/groups/slubsite/slubarchiv/SLUB-mediaconch-profile_2023.1.zip] (18.03) hinterlegt.

Glossar

Aufnahme

besteht aus einem oder mehreren Werken bzw. Takes, welche nacheinander auf einem Tonband aufgenommen wurden

Mehrspurverfahren

parallele Aufnahme mehrerer, voneinander unabhängiger Signale auf 2 … 64 Tonspuren (eine Aufnahme vorwärts, aber verschiedene Audiosignale je Spur, z.B. unterschiedliche Instrumente in Mono oder Stereo, die im Zusammenklang die Aufnahme bilden)

Rückwärtsbetrieb

gleichzeitige Abtastung einer zweiten Tonspur mit einem Mehrspurgerät, die sonst im Abspielbetrieb eine Laufrichtungsumkehr erfordert hätte, zB. Rückseite einer Tonbandkasette

Takes

Teilstück einer Aufnahme, die ohne Unterbrechung aufgezeichnet wurde

Tonspuren

Aufteilung eines Magnetbandes in mindestens zwei parallel verlaufende Aufzeichnungsbereiche

Bibliographie

Bennett, W. R. (1948). Spectra of quantized signals. *The Bell System Technical Journal*, *27*(3), 446–472. https://ieeexplore.ieee.org/abstract/document/6773026

Kester, W. (2022). *MT001: Taking the Mystery out of the Infamous Formula, "SNR=6.02N + 1.76dB," and Why You Should Care*.

Meyer, E. B., & Moran, D. R. (2007). Audibility of a CD-Standard A/DA/A Loop Inserted into High-Resolution Audio Playback. *Journal of the Audio Engineering Society*, *55*(9), 775–779.

Shannon, C. E. (1984). Communication in the presence of noise. *Proceedings of the IEEE*, *72*(9), 1192–1201. https://doi.org/10.1109/PROC.1984.12998

Unser, M. (2000). Sampling-50 years after Shannon. *Proceedings of the IEEE*, *88*(4), 569–587. https://doi.org/10.1109/5.843002

Webers, J. (1974). *Tonstudiotechnik 1974*. Deutsches Hifi-Museum. http://www.hifimuseum.de/2141.html

[1] FLAC ist zugleich die Bezeichnung für einen Codec, sowie für ein Dateiformat, welches diesen Codec nutzt.
[2] Hintergrund ist, dass diese Materialien andere signifikante Eigenschaften aufweisen
[3] Die A/D-Wandler können systembedingt intern andere Abtastfrequenzen und Quantisierungen verwenden, zB. im Falle von sigma-delta Abtastung. In dem Fall sind die Digitalisierungseinstellungen so zu wählen, dass sie einer sample-and-hold Abtastung mit 2^16 Quantisierungsstufen und der Abtastfrequenz von 48kHz (oder höher) entsprechen. Die in dieser Handreichung genannten Parameter beziehen sich auf das Digitalisat als Endergebnis der Digitalisierung.
[4] Der Faktor 2,1 ist den Vorgaben der AudioCD entlehnt und sollte für alle gängigen A/D-Wandler ausreichend Reserve zur Verfügung stellen.