

# Die Bewertung von Audioqualität in alltäglichen Hörsituationen

Michael Oehler<sup>1</sup>, Thomas Wildenburg<sup>1</sup>, Christoph Reuter<sup>2</sup>

<sup>1</sup> Institute for Sound and Vibration Engineering, University of Applied Sciences Düsseldorf

<sup>2</sup> Musicological Department, University of Vienna

## Hintergrund

Die Qualität von verlustbehaftet kodierten Audiodateien war und ist Gegenstand vieler Untersuchungen. Das Spektrum des experimentellen Fokus reicht u.a. von der grundsätzlichen Unterscheidung objektiver und subjektiver Messmethoden (Pocta & Beerends, 2015) über Teststandards (z.B. ITU-R BS.1116-1 oder ITU-R BS.1534; vgl. Breebaart, 2015) bis hin zu technischen Details der digitalen Signalverarbeitung (Tsai & Yen 2002; Zölzer, 2013; Khaldi et al. 2013). Kaum untersucht ist bisher jedoch die Relevanz der messbaren Qualitätsunterschiede in alltäglichen Hörsituationen, d.h. die Wahrnehmbarkeit von Kodierungsartefakten bei gleichzeitig vorhandenen Störgeräuschen.

Da ein immer größerer Anteil des Konsums populärer Musik in alltäglichen Hörsituationen stattfindet (vgl. BVMI Jahrbuch 2015), soll geklärt werden, wie sich diese auf die qualitative Bewertung verschiedener Kompressionsalgorithmen auswirken. Es wird vermutet, dass aufgrund von spektraler und zeitlicher Verdeckung bestimmte durch verlustbehaftete Kodierung entstehende ansonsten wahrnehmbare Artefakte nicht mehr erkannt werden (vgl. Lerch 2008).

## Ergebnisse

In einer Varianzanalyse mit Messwiederholung zeigte sich für die Qualitätsurteile bezüglich der verschiedenen Audioformate bzw. Bitraten sowohl für die Bedingung mit ( $F(10,20) = 25.237, p = .0001, \mu^2 = .927$ ) als auch ohne Störgeräusche ( $F(10,20) = 51.073, p = .0001, \mu^2 = .962$ ) ein signifikanter Unterschied. In beiden Bedingungen waren es genau die gleichen Audioformate, die im Vergleich zum Original signifikant unterschiedlich bewertet wurden (AAC 96, AAC 64, MP3 96, MP3 64, Vorbis 64). Die Effektstärken für die Varianten mit Störgeräuschen waren jedoch durchweg niedriger als für die Beispiele ohne Störgeräusche. Vergleicht man die beiden Bedingungen für jedes getestete Audioformat, so zeigt sich, dass signifikante Unterschiede nur für AAC 192 ( $F(1,29) = 20.054, p = .0001, \mu^2 = .409$ ), AAC 96 ( $F(1,29) = 14.483, p = .001, \mu^2 = .333$ ), AAC 64 ( $F(1,29) = 7.673, p = .01, \mu^2 = .209$ ) und Vorbis 192 ( $F(1,29) = 14.504, p = .001, \mu^2 = .333$ ) bestehen.

## Zusammenfassung

Es ist zu erkennen, dass bei den qualitativ besseren Klangbeispielen, d.h. generell bei höheren Bitraten sowie beim AAC- und Vorbis-Format z.T. auch bei geringeren Bitraten, die Beispiele mit Störgeräuschen signifikant schlechter bewertet werden. Bei den Aufnahmen mit qualitativ schlechterer Kodierung hingegen ist die alltägliche Hörsituation von geringerer Bedeutung für die Bewertung. Um die qualitativ besseren Kodierungen weiter differenzieren zu können, wird in einer bereits begonnenen Folgestudie anstatt des MUSHRA-Verfahrens die ABC/HR-Methode (ITU-R BS.1116) verwendet.

## Methode

Als Stimuli wurden Ausschnitte populärer Musik verwendet, bei welchen mögliche Kodierungsartefakte (Vorecho, Bandkompression, Verzerrung Räumlichkeit etc.) besonders deutlich wahrnehmbar sind. Die Kodierung der Ausschnitte geschah in Anlehnung an Audioformate, die häufig bei digitalen Musikdiensten zu finden sind (mp3, AAC sowie Vorbis mit 64, 96 und 192 kbps).

Die kodierten Dateien wurden zusammen mit der Originalaufnahme von 30 Vpn bewertet (MUSHRA-Test nach ITU-R BS.1534). Innerhalb eines within-subjects design bewerteten die Vpn sowohl eine Bedingung ohne Störgeräusche als auch eine alltägliche Bedingung mit Umgebungsgeräuschen. Diese wurden zuvor als Binauralaufnahme in einer belebten Fußgängerzone aufgezeichnet und später über Stax - SR-Lambda Pro Kopfhörer wiedergegeben. Die Wiedergabe der Musikausschnitte erfolgte in beiden Bedingungen über ein Samsung In Ear Stereo Headset EO-EG900BW (in der Bedingung mit Umgebungsgeräuschen unter dem Stax-Kopfhörer platziert).

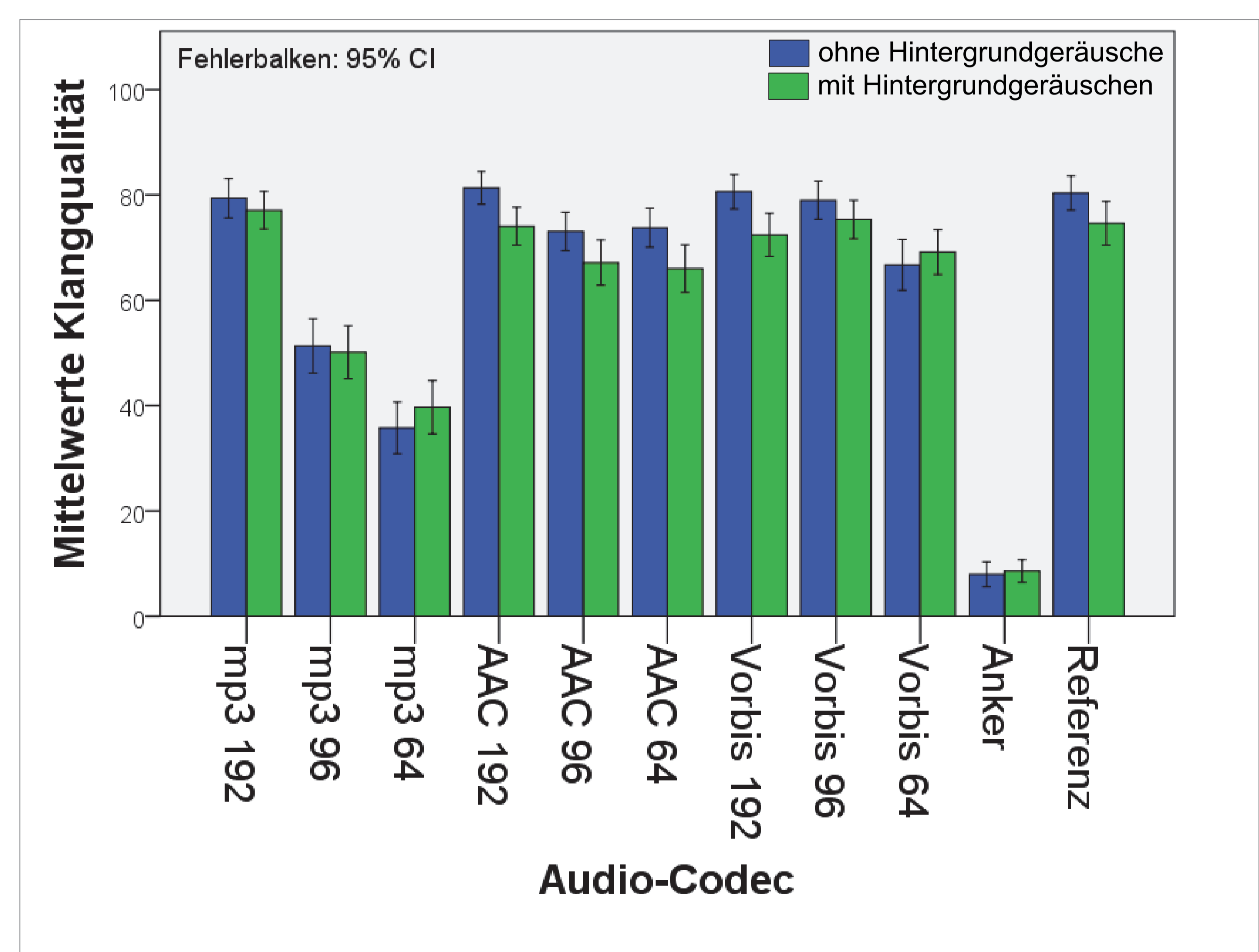


Abb. 1: Bewertung der verschiedenen Audio-Codec in einer alltäglichen Hörsituation (grün) sowie ohne Hintergrundgeräusche (blau).

## Literatur

Breebaart, J. (2015). Evaluation of statistical inference tests applied to subjective audio quality data with small sample size. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 23(5), 887-897.

BVMI (2015). „BVMI-2015-Jahrbuch-ePaper.pdf“.

ITU-R (1994). Recommendation BS. 1116: Methods for subjective assessment of small impairments in audio systems including multichannel sound systems, International Telecommunications Union.

ITU-R (2001). Recommendation BS. 1534: Method for the subjective assessment of intermediate quality level of coding systems, International Telecommunications Union.

Khaldi, K., Boudraa, A. O., Torresani, B., & Chonavel, T. (2015). HHT-based audio coding. *Signal, image and video processing*, 9(1), 107-115.

Lerch, A. (2008). Bitratenreduktion. In *Handbuch der Audiotechnik* (pp. 849-884). Springer Berlin Heidelberg.

Poča, P., & Beerends, J. G. (2015). Subjective and objective assessment of perceived audio quality of current digital audio broadcasting systems and web-casting applications. *IEEE Transactions on Broadcasting*, 61(3), 407-415.

Tsai, T. H., & Yen, C. C. (2002). A high quality re-quantization/quantization method for MP3 and MPEG-4 AAC audio coding. In *Circuits and Systems, 2002. ISCAS 2002. IEEE International Symposium on* (Vol. 3, pp. 851-854). IEEE.

Zölzer, U. (2013). *Digitale Audiosignalverarbeitung*. Springer-Verlag.