

Was macht Musik „hart“? Klangliche Merkmale zur genreübergreifenden Identifikation musikalischer Härte

Isabella Czedik-Eysenberg¹, Denis Knauf², Christoph Reuter¹

¹ *Institut für Musikwissenschaft, Universität Wien, Email: isabella.czedik-eysenberg@univie.ac.at*

² *Software & Information Engineering, Technische Universität Wien*

Hintergrund

Während in zahlreichen Untersuchungen sogenannte „harte“ Musik - meist Heavy Metal - in ihrer musiksoziologischen und -psychologischen Rolle durchleuchtet wird, etwa in Bezug zu Themen wie aggressivem Verhalten unter Jugendlichen [17], Persönlichkeit und Emotionsmodulation [19] und ihrem subkulturellen Umfeld [18][16][8], spielt hierbei oft nur am Rande die Frage mit, welche klanglichen Eigenschaften die im Mittelpunkt der Betrachtungen stehende Musik als „hart“ kennzeichnen.

In Artikeln zur Produktion entsprechender Musik und anderen beschreibenden Arbeiten findet sich eine Reihe charakterisierender Eigenschaften: Genannt werden mitunter etwa stark verzerrte Gitarrenklänge [16][1], eine ausgeprägte Intensität tiefer, bzw. auch hoher Frequenzbereiche [14][2][11], hohe Lautheit in Verbindung mit einem geringen Dynamikumfang [18][16], insbesondere eine flache dynamische Hüllkurve infolge von Klangverzerrungen [2], starke perkussive Klänge [8], ein hoher Geräuschanteil in Bezug auf das stimmliche Timbre [15], unklare Tonalität bzw. harmonische Dissonanzen [1] und ein besonders schnelles oder auch langsames Tempo [15].

Dabei wird „harte Musik“ besonders häufig mit dem Genre Metal assoziiert [14], kann aber auch eine Reihe weiterer Musikrichtungen wie etwa Hard Rock, Hardcore-Techno oder Neue Deutsche Härte umfassen. Insofern stellt sich auch die Frage, inwieweit sich für die Empfindung von „Härte“ allgemeine - genreübergreifende - Merkmale identifizieren lassen.

Fragestellungen und Ziele

- In Bezug auf welche akustischen Signaleigenschaften lässt sich ein wesentlicher Zusammenhang zu als „hart“ bewerteten Audio-Stimuli erkennen?
- In welchem Maße ist es möglich, eine automatisierte Detektion musikalischer Härte ausschließlich anhand jener Messwerte durchzuführen?
- Lassen sich übereinstimmende klangliche Merkmale zwischen Heavy Metal und anderen als „hart“ bezeichneten Musikgenres finden, und so ein allgemeines Messmodell für musikalische Härte aufstellen?
- Unterscheiden sich die Bewertungen von Hörern, welche „harte“ Musik mögen und/oder häufig konsumieren von jenen anderer Probanden, und wenn ja - inwiefern könnten hier andere akustische Kriterien relevant sein?

Methode

Hörversuch

Innerhalb einer mittels Ruby on Rails entwickelten Online-Versuchsumgebung [12] wurden 62 kurze Musikausschnitte unterschiedlicher Genres als Stimuli dargeboten. Dabei wurden die Versuchspersonen dazu aufgefordert, auf einer 10-stufigen Skala anzugeben, in welchem Maße sie vorgespielte Stimuli subjektiv als „hart“ wahrnehmen. In einem allgemeinen Fragebogen wurden zuvor neben deskriptiven Daten (Alter, Geschlecht) Informationen zur persönlichen Präferenz für harte Musik, sowie der Häufigkeit des Konsums abgefragt. Im Zuge der Bewertung gaben die Versuchspersonen auch an, ob der vorliegende Musikausschnitt beziehungsweise dessen Interpret ihnen bekannt ist. Abschließend wurden sie um eine verbale Beschreibung der ihrer Meinung nach für Wahrnehmung musikalischer Härte relevanten Kriterien gebeten.

Die Bewertung erfolgte durch 40 Versuchsteilnehmer im Alter von 18 bis 59 Jahren (Durchschnittsalter 31,08; 15 Frauen und 25 Männer).

Klanganalysen

Die verwendeten Klangbeispiele wurden Signalanalysen in MATLAB unter Einsatz der TSM Toolbox [4], der MIRtoolbox [9], sowie der Genesis Loudness Toolbox [7] unterzogen. Im Zuge dieser wurden insbesondere Niederfrequenzanteile, Inharmonizitäten, Rauigkeit, Lautheit, dynamischer Umfang und Hüllkurvenverlauf, Tempo sowie perkussive Anteile betrachtet. Anschließend wurde ermittelt, ob und auf welche Weise jene Eigenschaften in Zusammenhang mit den erhobenen Bewertungen durch die Versuchspersonen stehen.

Ergebnisse

Untersuchte Signaldeskriptoren

Die von den Versuchsteilnehmern gegebenen verbalen Beschreibungen wurden nach Synonymen zusammengefasst und hinsichtlich der Häufigkeit ihrer Nennung aufgereiht (siehe Tabelle 1).

Ausgehend davon, sowie von bestehenden Beschreibungen der Charakteristika von Metal und anderen als „hart“ angesehenen Musikrichtungen, wurden entsprechende Signalparameter als Pendant der psychoakustischen Eigenschaften extrahiert.

Tabelle 1: Von Versuchspersonen genannte Charakteristika für Härte

Genannte Eigenschaft	Nennungen
„hohes Tempo“	26
„spezieller Gesangsstil“	17
„wenig Melodie“	15
„hohe Lautheit“	14
„Schlagzeugpräsenz“	14
„dominanter Bass“	13
„Verzerrungen“	9
„spezifische Gitarrenriffs“	8
„Geräuschhaftigkeit“	6
„zeitliche Klangdichte“	6

In einer Korrelationsanalyse wurden die erlangten Deskriptorwerte den Durchschnittsbewertungen aller Musik-Stimuli gegenübergestellt. Tabelle 2 führt die stärksten Korrelationen an.

Als besonders relevant zeigte sich dabei etwa die Intensität perkussiver Signalkomponenten. Um diese zu extrahieren wurde das von Driedger und Müller implementierte Verfahren zur Harmonisch-Perkussiv-Zerlegung von Audiosignalen mittels Median-Filterung [4] nach Fitzgerald [6] eingesetzt. Je stärker der gemessene Effektivwert der verbleibenden perkussiven Signalkomponente nach erfolgter Zerlegung insgesamt war, desto höher fiel die Härtebewertung in der Regel aus. Diese starke Ausprägung perkussiver Signalkomponenten, welche sich in einem Sonagramm als vertikale Strukturen niederschlagen [5][6], ist kongruent mit Beschreibungen intensiver Schlagzeugklänge als ein zentrales Element harter Musik (etwa bei [14, S. 69], oder „pounding percussion“ bei [8, S. 119]). Dies schlägt sich auch in Form von spektralen Fluktuationen nieder (siehe Korrelation mit Spectral Flux (Median) in Tabelle 2).

Tabelle 2: Mit der Härtebewertung zusammenhängende akustische Signaleigenschaften

Klangdeskriptor	Korrelation mit Härte	
	r	p
Percussive Energy	0,81	< 0,01
Spectral Flux (Median)	0,80	< 0,01
Roughness	0,75	< 0,01
Number of Onsets	0,68	< 0,01
High Frequency Ratio	0,59	< 0,01
Loudness (Sone)	0,54	< 0,01
Low Centroid Rate	-0,52	< 0,01
2-4 kHz Energy	0,51	< 0,01
Envelope Flatness	0,50	< 0,01
Low Frequency Ratio	0,48	< 0,01
Inharmonicity	0,25	< 0,05

Die von Versuchspersonen besonders häufig angeführte Tempoabhängigkeit, und insbesondere auch hohe „zeitliche Klangdichte“, spiegelte sich in der Anzahl der Onsets wider. Dabei wurden mithilfe der MIRtoolbox [9] die

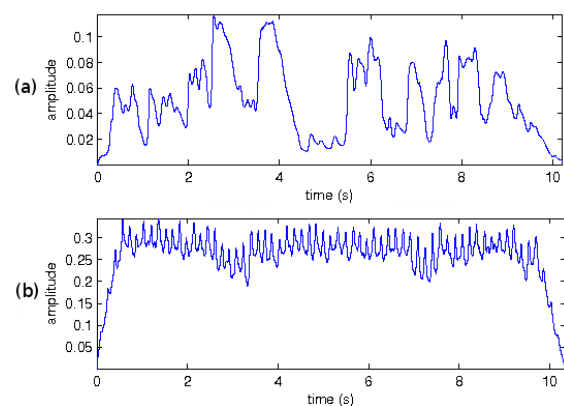
Einsätze anhand der zeitlichen Hüllkurve abgeschätzt. Deren Anzahl im Verhältnis zur Länge des Signals wurde bestimmt, wobei Musikbeispiele im Allgemeinen als umso härter bewertet wurden, je dichtere Onsets vorhanden waren.

Auch bestätigen lässt sich ein Zusammenhang der Härtebewertung mit einer intensiveren Ausprägung hoher (High Frequency Ratio) so wie tiefer Frequenzbereiche (Low Frequency Ratio) gegenüber den Mittenfrequenzen. Hierbei wurde mittels Bandpassfiltern der Bereich unterhalb von 100 Hz beziehungsweise oberhalb von 1000 Hz extrahiert und hinsichtlich des Effektivwertes mit dem Frequenzband 250 bis 400 Hz verglichen. Dieser Zusammenhang steht im Einklang mit den Ergebnissen von Berger und Fales, welche zeigten, dass Gitarrentimbres mit zunehmendem Gehalt an hochfrequenter Energie als härter wahrgenommen werden [2, S. 123-194].

Hervorzuheben ist dabei auch die Korrelation mit der Intensität der Signalenergie im Bereich zwischen 2 und 4 kHz. Dieses Frequenzband kann durch seine Lage im Bereich der Eigenresonanzen des Außenohres dazu beitragen, dass ein Klang - etwa Tafelkratzen - als unangenehm wahrgenommen wird [13].

Für die Bestimmung der Lautheit (Sone) wurde die in der Genesis Loudness Toolbox implementierte Messung gemäß der ANSI-2007-Norm nach Moore et al. [10] zu Rate gezogen. Auch die oft genannte höhere Lautheit harter Musik [16][18] lässt sich somit feststellen.

Detaillierter zeigt sich das auch anhand der dynamischen Hüllkurve typischer Beispiele. Betrachtet man diese etwa im Fall des am härtesten bewerteten Musikbeispiels „Marduk - Slay the Nazarene“ (Abbildung 1 (b)) im Vergleich zu dem Musikstück mit der niedrigsten Härtebewertung „Cat Stevens - Sad Lisa“ (Abbildung 1 (a)), fällt nicht nur ein höherer Maximalpegel auf, sondern auch eine im Verhältnis insgesamt flachere Hüllkurve, geprägt von zeitlich dicht aufeinanderfolgenden Spitzen mit relativ geringen Absenkungen dazwischen.

**Abbildung 1:** Dynamische Hüllkurve des (a) am wenigsten hart bewerteten Beispiels „Cat Stevens - Sad Lisa“ und des (b) am härtesten bewerteten Beispiels „Marduk - Slay the Nazarene“

Jene konstante (perkussive) Ereignisdichte und die da-

mit einhergehenden gleichmäßigeren Signaleigenschaften spiegeln sich in Korrelationen mit den Messwerten Number of Onsets, Envelope Flatness, aber auch Low Centroid Rate (der Anteil aller Frames, deren Spektraler Schwerpunkt unterhalb des Durchschnittes liegt) wider.

Insgesamt ist jedoch zu berücksichtigen, dass einige der untersuchten Signaleigenschaften stark untereinander zusammenhängen (z.B. Percussive Energy und Spectral Flux $r = 0,98$; $p < 0,01$). Daher wurde mittels einer Analyse partieller Korrelationen eruiert, welche Eigenschaften in Kombination unabhängige Beiträge zu der Erklärung der Härteempfindung leisten können und so geeignete Sets an Dimensionen ermittelt.

Von allen erhobenen Merkmalen zeigte sich die Kombination aus Perkussiver Energie und Gesangsstil (nach manueller Klassifikation) als besonders effizient zur Beschreibung der insgesamten Härteempfindung (siehe Abbildung 2). Ein entsprechendes mittels multipler linearer Regression aufgestelltes Modell korreliert stark mit den Bewertungen der Versuchspersonen ($r = 0,88$; $p < 0,01$).

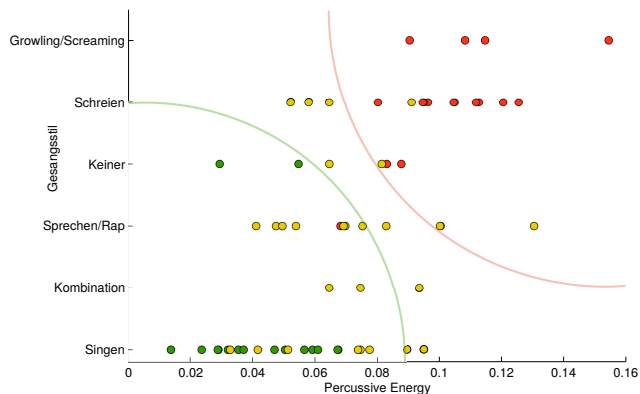


Abbildung 2: Klangbeispiele nach Intensität perkussiver Signalkomponenten und Gesangsstil. Rote Datensätze liegen im obersten Viertel hinsichtlich der Härtebewertung, grüne im untersten.

Werden manuell klassifizierte Merkmale ausgespart und nur von der Menge der automatisiert extrahierten Signaleigenschaften ausgegangen, erweist sich hier die Kombination aus perkussiver Energie (E_{perc}), Intensität der Signalkomponenten zwischen 2 und 4 kHz (E_{2-4kHz}) und Low Centroid Rate (LCR) als die effektivste (siehe Abbildung 3).

Mittels multipler linearer Regression unter Einbeziehung der mittleren Härtebewertung als Zielgröße wurde ein automatisiert extrahierbarer Härte-Deskriptor gemäß (1) aufgestellt, welcher eine starke Übereinstimmung mit den Härtebewertungen zeigt ($r = 0,86$; $p < 0,01$; siehe Abbildung 4).

$$\begin{aligned}
 \text{Härte} \approx & 2,67 + 33,80 \cdot E_{perc} \\
 & + 6,37 \cdot E_{2-4kHz} \\
 & - 4,65 \cdot LCR
 \end{aligned} \tag{1}$$

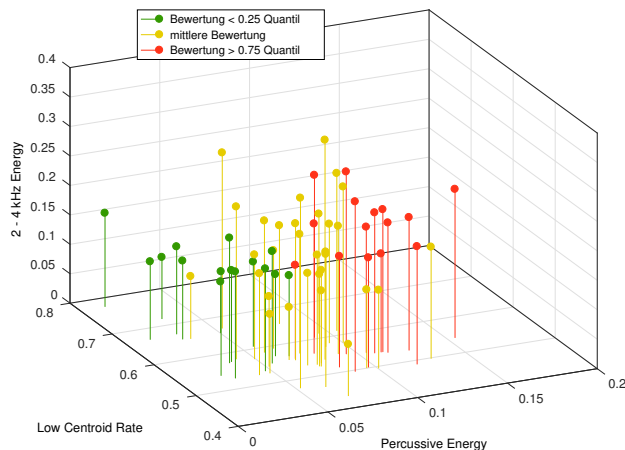


Abbildung 3: Die drei Eigenschaften Percussive Energy, Low Centroid Rate und 2-4 kHz Energy weisen in Kombination einen starken Zusammenhang mit den Härteurteilen der Versuchspersonen auf.

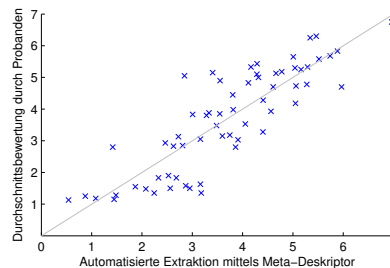


Abbildung 4: Korrelation zwischen aufgestelltem Härte-Modell und Bewertungen im Hörversuch ($r = 0,86$, $p < 0,01$)

Einfluss von subjektiven Präferenzen und Musikgenre

Probanden mit positiver Präferenz gegenüber harter Musik waren signifikant häufiger männlich ($p < 0,05$ laut exaktem Fisher-Test) und wiesen ein niedrigeres Alter auf ($p < 0,01$ laut t-Test). Jene „Hart-Hörer“ bewerteten Stimuli im Durchschnitt signifikant als etwas weniger hart als jene mit einer negativen Präferenz ($p < 0,05$ laut t-Test). Hörer mit negativer Präferenz gegenüber harter Musik beurteilten Beispiele aus dem Bereich Techno/Hardcore signifikant als härter als Hörer harter Musik ($p < 0,01$ laut t-Test), während das etwa auf Black/Death Metal nicht zutrif.

Die Härtebewertungen variierten mit dem Genre der Musikausschnitte, wobei Beispiele aus den Bereichen Black/Death Metal, Techno/Hardcore, Metal und Neue Deutsche Härte im Schnitt als am härtesten bewertet wurden. Für verschiedene betrachtete Genres zeigten sich im Wesentlichen übereinstimmende akustische Kriterien zur Bildung eines Härte-Metadeskriptors. Eine Ausnahme bildete hierbei jedoch Techno/Hardcore, bei dem der Flachheit der Hüllkurve alleine eine deutlich größere Rolle zuzukommen scheint (Korrelation: $r = 0,90$; $p < 0,01$).

Zusammenfassung und Ausblick

Verschiedene beschriebene Merkmale musikalischer Härte konnten in Form psychoakustischer Signaleigenschaften quantifiziert und statistisch bestätigt werden. Dabei zeigten sich im Wesentlichen genreübergreifende Kriterien, insbesondere:

- Hohe Präsenz perkussiver Instrumente, die sich in Form spektraler Fluktuationen und hoher Intensität entsprechender Signalanteile niederschlägt (Percussive Energy, Spectral Flux)
- Spezieller Gesangsstil (Taschenfaltenstimme, Screaming)
- Flache Hüllkurve in Verbindung mit einem zeitlich dichten Klangteppich (Envelope Flatness, Number of Onsets)
- Hohe Lautheit, mit ausgeprägter Intensität insbesondere tiefer und hochfrequenter Spektralbereiche (Loudness (Sone), Low / High Frequency Ratio, 2-4 kHz Energy)

Zu bedenken bleibt, dass außerakustische Kriterien wie musikalische Metadaten, textuelle Inhalte oder auch etwa kompositorische Charakteristika wie spezifische Metal-Riffs hierbei nicht berücksichtigt werden und als semantische Lücke bestehen bleiben. Dennoch konnte anhand signalanalytischer Kriterien ein automatisiert extrahierbarer Gesamt-Deskriptor mit einer hohen Übereinstimmung mit den subjektiven Bewertungen der Versuchspersonen aufgestellt werden ($r = 0,86$; $p < 0.01$). In zukünftigen Untersuchungen soll die Eignung des Modells zu einer Vorhersage der Härte-Empfindung noch eingehender evaluiert werden, mit dem Ziel, es in ein System zur Extraktion psychoakustischer High-Level-Eigenschaften [3] zu integrieren.

Literatur

- [1] Berger, H.: *Metal, Rock and Jazz: Perception and the Phenomenology of Musical Experience*. Hanover, N.H.: Wesleyan University Press/University Press of New England, 1999
- [2] Berger, H. & Fales, C.: *Heaviness in the Perception of Heavy Metal Guitar Timbres: The Match of Perceptual and Acoustic Features over Time*. *Wired for Sound: Engineering and Technologies in Sonic Cultures*. Middletown, CT: Wesleyan University Press, 2005
- [3] Czedik-Eysenberg, I.: *Music Information Retrieval und Klangfarbe. „TimbRet“: Ein interaktives System zur inhaltsbasierten Suche komplexer Klänge anhand von High-Level-Deskriptoren*, Masterarbeit, Universität Wien, 2016
- [4] Driedger, J. & Müller, M.: TSM Toolbox: MATLAB Implementations of Time-Scale Modification Algorithms. *International Conference on Digital Audio Effects*, September 2014, 249-256
- [5] Driedger, J., & Müller, M.: Harmonisch-Perkussiv-Rest Zerlegung von Musiksignalen. *Proceedings of the Deutsche Jahrestagung für Akustik (DAGA)*, 2015, 1421-1424
- [6] Fitzgerald, D.: Harmonic/percussive separation using median filtering. *Proceedings of the International Conference on Digital Audio Effects (DAFX)*, Graz, Austria, 2010, 246-253
- [7] Genesis Loudness Toolbox, URL: http://genesis-acoustics.com/en/loudness_online-32.html
- [8] Gross, R. L.: Heavy metal music: A new subculture in American society. *The Journal of Popular Culture*, 24(1) (1990), 119-130
- [9] Lartillot, O. & Toivianen, P.: A Matlab toolbox for musical feature extraction from audio. *International Conference on Digital Audio Effects*, September 2007, 237-244
- [10] Moore, B. C., Glasberg, B. R., & Baer, T.: A model for the prediction of thresholds, loudness, and partial loudness. *Journal of the Audio Engineering Society*, 45(4) (1997), 224-240
- [11] Mynett, M.: Achieving intelligibility whilst maintaining heaviness when producing contemporary metal music. *Journal on the Art of Record Production* 6 (2012)
- [12] Protrabant Online-Versuchsumgebung, URL: <https://denkn.at/protrabant>
- [13] Reuter, C., Oehler, M. & Mühlhans, J.: Physiological and acoustical correlates of unpleasant sounds, *Proceedings of the Joint Conference ICMPC13-APSCOM5*, August 4-8, 2014, Yonsei University, Seoul, Korea, 97
- [14] Reyes, I.: *Sound, Technology, and interpretation in Subcultures of Heavy Music Production*. Dissertation - Pittsburgh University, 2008
- [15] Wallach, J., Berger, H. M. & Greene, P. D.: *Metal rules the globe: heavy metal music around the world*. Duke University Press, 2011, 180ff
- [16] Walser, R.: *Running with the devil: Power, gender, and madness in heavy metal music*. Wesleyan University Press, 1993
- [17] Weindl, D.: *Musik & Aggression. Untersucht anhand des Musikgenres Heavy Metal*. Peter Lang, Frankfurt am Main, 2005
- [18] Weinstein, D.: *Heavy Metal: A Cultural Sociology*. New York, N.Y.: Maxwell Macmillan International, 1991
- [19] von Georgi, R., Kraus, H., Cimbal, K., & Schütz, M.: Persönlichkeit und Emotionsmodulation mittels Musik bei Heavy-Metal Fans. *Musikpsychologie. Jahrbuch der Deutschen Gesellschaft für Musikpsychologie*, Bd. 21 (2011), 90-118