Formanten bringen Farbe ins Spiel
Möglichkeiten der Klangfarbenbeschreibung im dynamischen Formantenfeld

Saleh Siddiq[1], Christoph Reuter[1], Isabella Czedik-Eysenberg[1], Michael Oehler[2]

[1] *Universität Wien, 1090 Wien, E-mail: christoph.reuter@univie.ac.at*

[2] *Universität Osnabrück, 49074 Osnabrück, michael.oehler@uni-osnabrueck.de*

Populäre Methoden der Klangfarben-beschreibung

In den letzten 50 Jahren konnten sich besonders drei Arten der Klangfarbenbeschreibung etablieren:

Timbre Spaces

Sieht man von den frühen Entwürfen zur räumlichen Darstellung von Klangfarben ab [1], gilt der 1975 von John Grey entwickelte Timbre Space als frühester Klangfarbenähnlichkeitsraum [2], in dem die Ähnlichkeiten von Klangfarben entlang dreier Raumachsen visualisiert werden (je näher die Klänge räumlich beieinanderliegen, desto ähnlicher klingend werden sie empfunden):

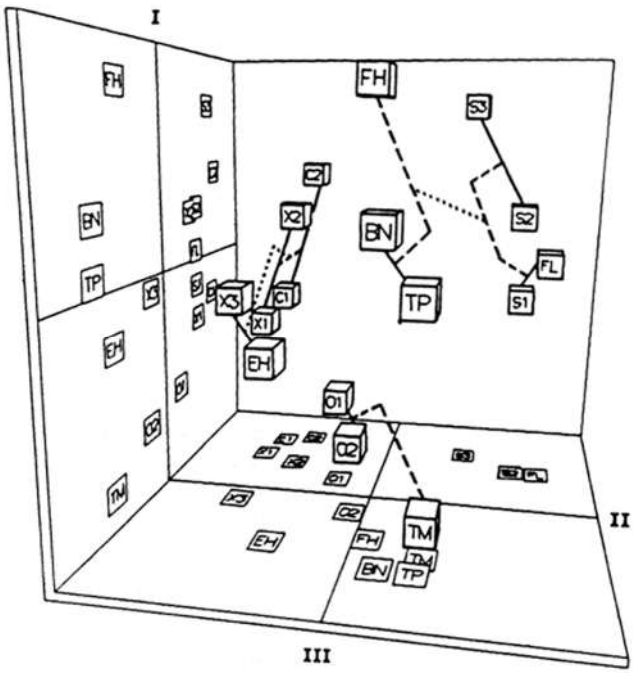

Abbildung 1: Timbre Space mit den Achsen I: spektrale Energieverteilung, II: Einschwingvorgang und Synchronizität der höheren Teiltöne dort, III: Fluktuationen und Inharmonizität [2].

Timbre Spaces sind zwar intuitiv, jedoch beruhen sie zumeist auf nur einer Tonhöhe pro (meist auch noch synthetisiertem) Instrument. Verschiedene Dynamikstufen und Artikulationsweisen werden hier ebenfalls nicht berücksichtigt. Ein Vergleich der drei bisher populärsten Timbre Spaces in einem Meta Timbre Space ergab zudem, dass Klangfarben desselben Stimuli-Sets untereinander ähnlicher waren als die Klangfarben desselben Instruments. D.h. Timbre Spaces sind zur Klangfarbenbeschreibung nur bedingt geeignet, da sie weder untereinander vergleichbar noch verallgemeinerbar sind [3].

Mel Frequency Cepstral Coefficients (MFCC)

Die von Davis und Mermelstein in den 1980er Jahren ursprünglich für die Erkennung von Sprecher*innen und Sprachähnlichkeiten entwickelten Mel Frequency Cepstral Coefficients (MFCC) [4] haben sich hingegen auch für die Ähnlichkeitsberechnung von Instrumentalklangfarben besonders gut bewährt. Auch wenn sie als transformierte Repräsentation der Spektralinformation weitaus weniger intuitiv als die Timbre Spaces sind, haben Sie sich in den letzten Jahren als robuste Standard-Features für die Berechnung von Klangfarbenähnlichkeiten etabliert.

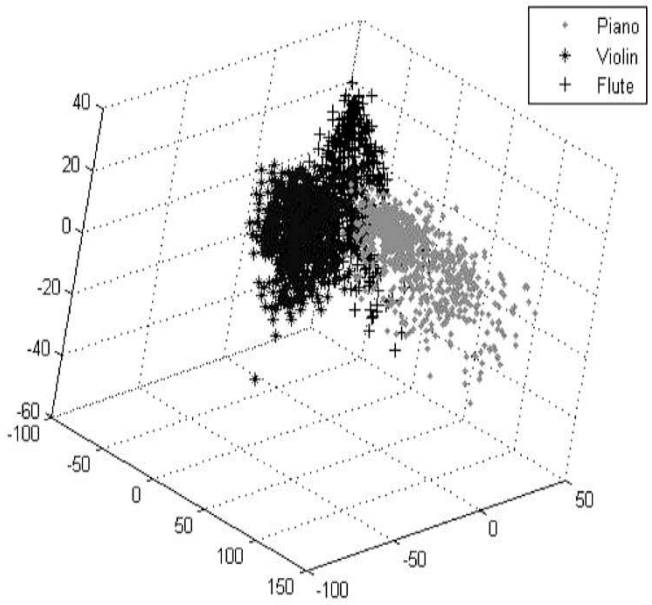

Abbildung 2: Automatische Klangfarbentrennung im MFCC-Space [5].

Formanten

In der deutschsprachigen Literatur sind darüber hinaus seit 1929 die von Karl Erich Schumann für Instrumentalklangfarben nachgewiesenen Formantbereiche zu finden [6], die als tonhöhenunabhängige stabile spektrale Maxima die Klangfarben vor allem von Blech- und Doppelrohrblattinstrumenten charakteristisch in ihrer jeweiligen Vokalähnlichkeit prägen. Da sie durch ihre Vokalähnlichkeit hörbar und im Spektrum direkt sichtbar sind, erlauben Sie eine intuitivere Einschätzung der Klangfarbe als die MFCCs.
Bei formantgeprägten Instrumenten haben sich Formanten zur Beschreibung von Klangfarbenähnlichkeiten als ähnlich zuverlässig herausgestellt wie MFCCs [7]. Sie haben sich sowohl in Experimenten zur Auditory Scene Analysis [8] als auch zur klanglichen Verschmelzung [9][10] bewährt und entsprechen überdies den in Instrumentationslehren aufgestellten Empfehlungen zur Mischung von Instrumentalklangfarben [11].

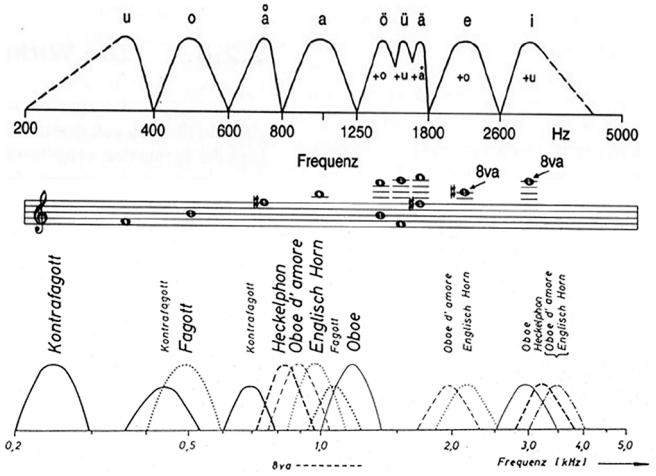

Abbildung 3: Vokalformanten und ihre Frequenzbereiche (oben), Formanten von Doppelrohrblattinstrumenten (unten)[12]

Formantberechnung in Libraries zur Signalanalyse musikalischer Klänge

Während die MFCCs direkt von der Phonetik in die musikalische Signalanalyse übernommen wurden, fehlen Berechnungsmöglichkeiten für die weitaus früher aus der Phonetik stammenden Formanten in den gängigen Libraries zur musikalischen Signalanalyse bislang völlig (z.B. in MIRToolbox [13], Timbre Toolbox [14], Essentia [15], Yaafe [16], Meyda [17], LibROSA [18], JSXtract [19], Aubio [20], MiningSuite [21]). Die Ursachen dafür liegen wahrscheinlich in dem Umstand, dass die grundlegenden Werke zu Formanten bei Musikinstrumenten [6][22] nicht ins Englische übersetzt wurden, so dass der Formantbegriff für Instrumentalklangfarben bis auf wenige Ausnahmen [10] [23][24] nur im deutschsprachigen Bereich Verbreitung fand. Dass die Formanten zur Beschreibung von Instrumentalklangfarben quasi „zu früh“ entdeckt wurden (d.h. lange bevor es überhaupt eine Idee von berechenbaren Timbre Features gab), mag mit ein Grund dafür sein, dass sie heute in Signalanalyse-Libraries zur musikalischen Klangfarbe fehlen. Man kann darüber hinaus natürlich anführen, dass die gängige Formantberechnung via LPC stark von den jeweiligen Voreinstellungen abhängig ist; dies ist jedoch auch bei der Helligkeits- oder Rauhigkeitsberechnung so, ohne dass deswegen auf diese Timbre Features verzichtet wird. So ist man heute für die Berechnung von Instrumentalformanten auf die Algorithmen aus eher auf Anwendungen innerhalb der Phonetik spezialisierten Audiosignalverarbeitungs-Libraries angewiesen (wie z.B. Colea [25], Praat [26], Parselmouth [27], Formant Tracker [28], Formant Estimation [29]). Diese schon bestehenden Algorithmen könnten mit einem geringen Aufwand in die Libraries und Toolboxen zur Signalanalyse von musikalischen Klängen übernommen werden.

Formantenfelder

Die Stärken der Formanten werden vor allem dann offenbar, wenn sie zur Klangfarbenbeschreibung in einem Formantenfeld eingesetzt werden. Ähnlich wie beim Vokaltrapez (und auf dieses beziehbar) werden hier die Instrumentalklänge anhand ihres ersten und zweiten Formanten entlang zweier Achsen angeordnet. Je näher die Klänge in diesem Feld beieinanderliegen, als desto ähnlicher klingend werden sie meist empfunden. Fallen sie dabei in einen bestimmten Vokalbereich, so lassen sie sich meist sehr gut durch diesen beschreiben (wie z.B. „Das Fagott klingt nach einem /o/“ oder „Die Trompete klingt nach einem /ä/“).

Ein erstes Formantenfeld für Musikinstrumentenklänge wurde 1974 von Udo Sirker als „Wahrnehmungsraum bei zweiformantigen Klängen“ [30] für Doppelrohrblattinstrumentenklänge erstellt. Die Position der Klänge in diesem Feld und ihre Zuordnung zu den jeweiligen Vokalen wurde aus den jeweils stärksten Amplituden im jeweiligen Spektrum ermittelt.

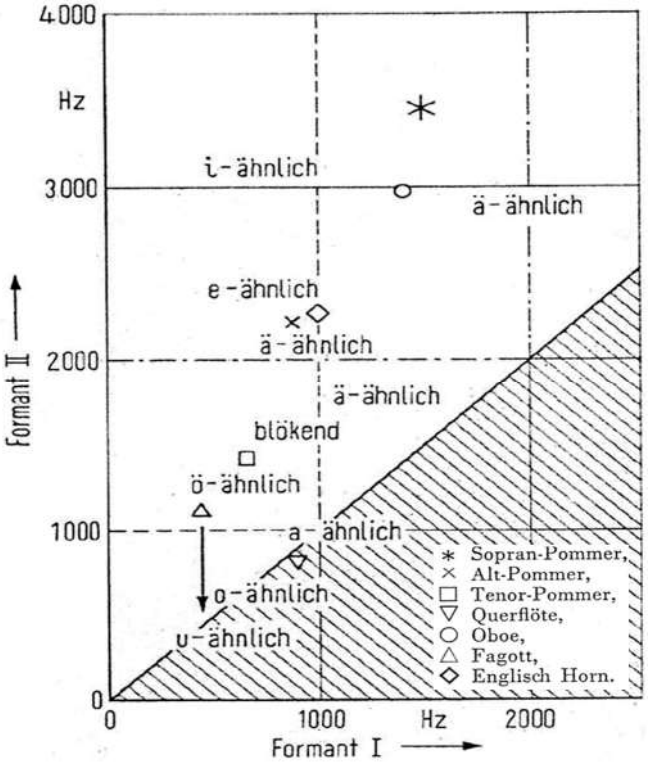

Abbildung 4: Doppelrohrblattinstrumentenklänge und ihre Vokalähnlichkeit im „Wahrnehmungsraum bei zweiformantigen Klängen“ [30].

Ein nächstes Formantenfeld folgte erst knapp 30 Jahre später, als John McCarty von der Stanford University die mit dem Matlab-Script Colea [25] ermittelten Formanten der gängigen Orchesterinstrumente auf ein Vokaltrapez projizierte (Abbildung 5). Auch wenn er pro Instrument nur eine Tonhöhe erhob (die auch noch von Instrument zu Instrument verschieden war) entdeckte auch er die Vorteile der Formantanalyse für Musikinstrumente: „Now I can say that the tuba's timbre has a 'U' sound or (oo) as in the word who, or the trumpet has an 'ae' timbre like the work actor. “ [24].

2016 folgte ein erstes interaktives Formantenfeld für die gängigen Holz- und Blechblasinstrumente, in welchem 593 Einzelklänge der Vienna Symphonic Library (VSL) in allen erreichbaren Tonhöhen, drei Registern und zwei Dynamikstufen (***ff*** und ***pp***) pro Instrument auf ein Vokaltrapez bezogen dargestellt werden ([31], Abbildung 6). Pro Instrument, Dynamikstufe und Register können hier auch die Mittelwerte (Kreismittelpunkt) und Standardabweichung (Durchmesser der Ellipsen) pro Formantbereich dargestellt werden.

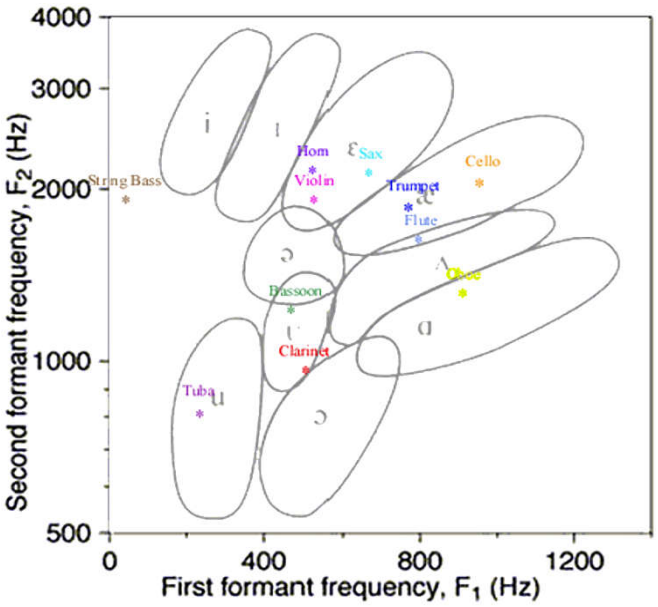

Abbildung 5: Klänge von Saxophon, Flöte, Oboe, Klarinette, Fagott, Horn, Trompete, Tuba, Geige, Cello, Bass in McCartys „Vowel Space“ [24].

Wahrnehmungsexperimente zur klanglichen Ähnlichkeit zeigten, dass im Formantenfeld (Abbildung 6) überlappende Bereiche als ähnlich klingend wahrgenommen werden, während weit auseinanderliegende Bereiche als unähnlich empfunden werden [32]. Darüber hinaus können die Formanten hier ähnlich effektiv zur Bestimmung von Klangähnlichkeiten eingesetzt werden wie MFCCs [7].

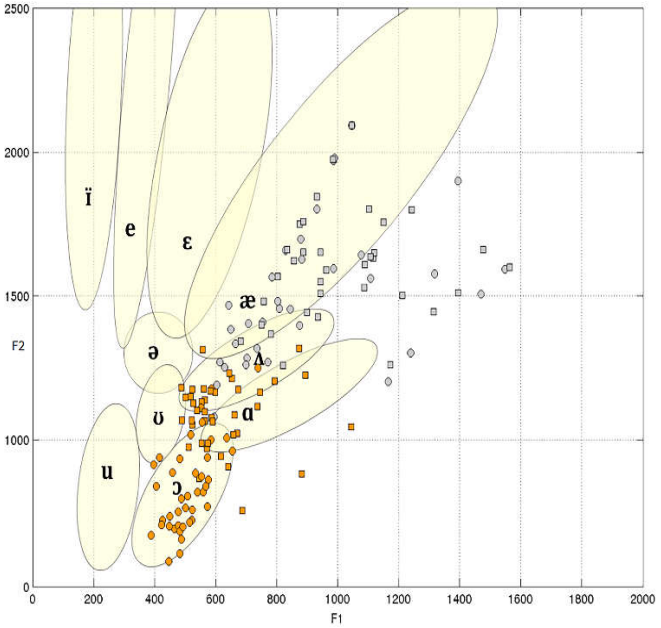

Abbildung 6: Formantenfeld mit den Klängen von Fagott (orange) und Oboe (grau) in allen erreichbaren Tonhöhen in ***ff*** (Quadrate) und ***pp*** (Kreise)[32].

2020 erfolgte ein Update des Formantenfelds von Flash auf Javascript unter [33], bei dem die Auswahl der Musikinstrumente um Geige, Viola, Cello und Bass erweitert wurde. Mit der Einbindung einer zusätzlichen Sample-Library für Blasinstrumente von Spitfire Audio erweiterte sich die Datenbasis auf 1100 ausgewertete Einzelklänge.

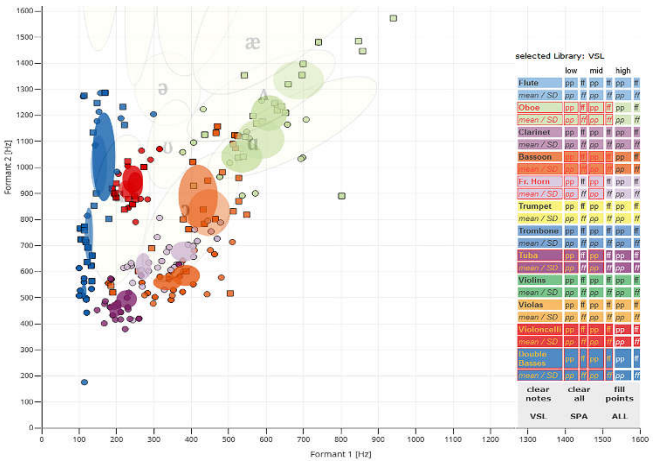

Abbildung 7: Formanten, ihre Mittelwerte und Standardabweichungen von Oboe (grün), Fagott (orange), Horn (rosa), Tuba (lila), Cello (rot) und Bass (blau) im tiefen und mittleren Register [33]

Dynamisches Formantenfeld

Während das Formantkonzept bislang nur an Einzelklängen geprüft wurde, können im dynamischen Formantenfeld unter [34] Formantbewegungen direkt im musikalischen Kontext dargestellt werden. Hierfür wurden via Parselmouth [27] die Formanten der einzelnen Musikinstrumente getrackt, die im 1. Satz der 7. Sinfonie Beethovens ihren Einsatz finden. Die hierfür verwendeten nachhallfreien Einzelspuren stammen aus [35]. Via Plotly und P5 wurden die erhobenen Werte mit den Audiodateien synchronisiert und auf ein interaktives Formantenfeld übertragen, so dass dadurch das Verhalten von Instrumentalformanten in „freier Wildbahn“ visualisierbar wird, die sich während des Stücks in jeweils spezifischen, für sie typischen „Revieren“ oder Bereichen bewegen.

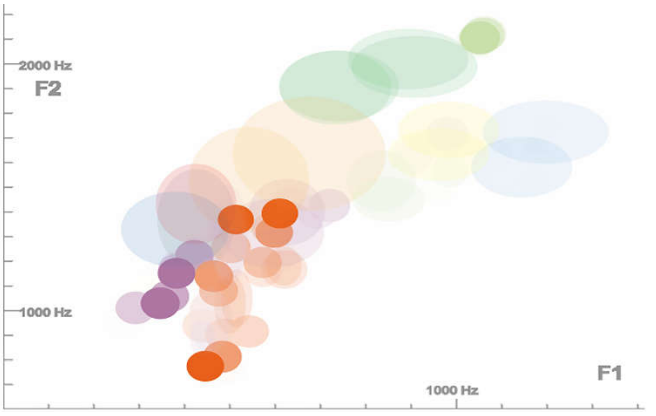

Abbildung 7: Formant-Mittelwerte und -Standardabweichungen der beteiligten Orchesterinstrumente (farbige Ellipsen); hier im Zusammenspiel: Hörner (lila), Fagotte (orange) und Oboe (grün) [34].

Ein ähnliches interaktives dynamisches Feld (mit einem vergleichbaren Verhalten der Klangfarben) lässt sich auch für den MFCC1 und MFCC2 erstellen [34], ebenso wie für zwei Dimensionen des Timbre Spaces (Spectral Centroid und Spectral Flux), bei denen es jedoch schwieriger ist typische klangliche Bereiche der Musikinstrumente zu erkennen [34]. Unter [34] können diese interaktiven dynamischen Klangfarbenfelder vergleichend ausprobiert werden. Über die

farbigen Schalter am oberen Bildrand des Feldes können jeweils beliebige Instrumentenkonfigurationen zusammengestellt werden bzw. Instrumente stumm- oder dazugeschaltet werden. Mit dem Button „all“ (ganz rechts in der oberen Leiste) lassen sich alle Instrumente während des Spiels ein- und ausschalten, während mit dem Button „Play“ das Feld gestartet werden kann.

Fazit und Zusammenfassung

Obwohl sowohl MFCCs als auch Formanten zur Beschreibung von klanglichen Eigenschaften aus der Phonetik in die Musikalische Akustik übernommen wurden, gibt es in den gängigen Musik-Signalanalyse-Libraries bislang nur die MFCCs. Dabei könnten die derzeit in Python- und Matlab-Scripten vorhandenen Formant-Extraktions-Algorithmen ohne weiteres auch zur Feature-Extraktion in die Musik/Klang-Signalanalyse-Libraries übernommen werden.

Spätestens seit den 1970er Jahren gibt es Ansätze dafür Musikinstrumente auf das Vokaltrapez der Phonetik bzw. das darauf basierende Formantenfeld abzubilden. Mit einer ausreichend hohen Anzahl an Einzelklängen in allen erreichbaren Tonhöhen und verschiedenen Dynamikstufen ergeben sich pro Register klare Bereiche, in denen sich die Instrumentalklangfarben im Formantenfeld ansiedeln.

Die im dynamischen Formantenfeld getrackten Formanten der Instrumente ermöglichen eine einfache und intuitive Visualisierung des Formantverhaltens bei Tonhöhen- und Dynamikwechseln und sind hilfreich bei der Darstellung von Klangfarbenähnlichkeiten und -unterschieden.

Literatur

[1] Albersheim, G.: Zur Psychologie der Ton- und Klangeigenschaften. Strassburg: Heitz 1939.

[2] Grey, J.: An Exploration of musical timbre. CCRMA Stanford University 1975.

[3] Siddiq, S. et al.: Towards the comparability and generality of timbre space studies. 3rd Vienna Talk on Music Acoustics 2015, S. 232–235.

[4] Davis, S.B., et al.: Comparison of parametric Representations for monosyllabic Word Recognition. IEEE Transactions on ASSP 28, 1980, S. 357–366.

[5] Loughran, R. et al.: Musical instrument identification using principal component analysis and multi-layered perceptrons. ICALIP 2008, S. 643–648.

[6] Schumann. K.E.: Physik der Klangfarben. Berlin 1929.

[7] Reuter, C. et al.: Formant Distances and the Similarity Perception of Wind Instrument Timbres. ICMPC15/ESCOM10, 2018, S. 367–371.

[8] Reuter, C.: Verschmelzung und partielle Verdeckung. 26. DAGA, Kiel 2000, S. 176–177.

[9] Reuter, C.: Die auditive Diskrimination von Orchesterinstrumenten. Frankfurt: Lang 1996.

[10] Lembke, S.-A. et al.: The Role of Spectral-Envelope Characteristics in perceptual Blending of Wind-Instrument Sounds. Acta Acustica 101, 2015, S. 1039–1051.

[11] Reuter, C.: Klangfarbe und Instrumentation. Frankfurt: Lang 2002.

[12] Meyer, J.: Akustik und musikalische Aufführungspraxis. Bergkirchen: PPV Medien 2015.

[13] Lartillot, O. et al.: A Matlab Toolbox for Musical Feature Extraction from Audio. 10th DAFx, Bordeaux 2007, S. 237–244.

[14] Peeters, G. et al.: The Timbre Toolbox. JASA, 130(5), 2011, S. 2902–2916.

[15] Bogdanov, D. et al.: ESSENTIA: an Audio Analysis Library for Music Information Retrieval. ISMIR'13, 2013, S. 493–498.

[16] Yaafe 2013, URL: https://github.com/Yaafe/Yaafe

[17] Meyda 2014, URL: https://meyda.js.org/

[18] LibROSA 2015, URL: https://librosa.org/

[19] JSXtract 2016, URL: https://github.com/nickjillings/js-xtract

[20] Aubio 2017, URL: https://aubio.org/

[21] Lartillot, O.: MiningSuite 2019, URL: http://olivierlar.github.io/miningsuite/

[22] Stumpf, C.: Die Sprachlaute. Berlin: Springer 1926.

[23] Luce, D.A. et al.: Physical Correlates of Brass-Instrument Tones. JASA 42(6), 1967, S. 1232–1243.

[24] McCarty, J.: Formant Analysis 2003, URL: https://ccrma.stanford.edu/%7Ejmccarty/formant.htm

[25] Loizou, P.: COLEA 1998, URL: https://ecs.utdallas.edu/loizou/speech/colea.htm

[26] Boersma, P., & Weenink, D. (2013). *Praat: Doing Phonetics by Computer*, Version 5.3.51, retrieved 2 June 2013, URL: http://www.praat.org/

[27] Jadoul, Y. et al.: Introducing Parselmouth: A Python interface to Praat. J. of Phonetics 71, 2018, S. 1–15.

[28] Kamth, N.: Formant Tracker, 2021, URL: https://de.mathworks.com/matlabcentral/fileexchange/8959-formant-tracker

[29] Rabiner, L. et al.: Formant Estimation, 2021, URL: https://de.mathworks.com/matlabcentral/fileexchange/45315-formant-estimation

[30] Sirker, U.: Strukturelle Gesetzmäßigkeiten in den Spektren von Blasinstrumentenklängen. Acustica 30, 1974, S. 49–59.

[31] Reuter, C.: Interaktives Formantenfeld 1.0, 2016, URL: https://muwiserver.synology.me/formanten/formantregister4_english.htm

[32] Reuter, C. et al.: Formanten als hilfreiche Timbre-Deskriptoren. 43. DAGA, Kiel 2017, S. 190–193.

[33] Reuter, C.: Interaktives Formantenfeld 2.0, 2020, URL: https://muwiserver.synology.me/formantmap

[34] Reuter, C.: Dynamisches Formantenfeld, 2021, URL: https://muwiserver.synology.me/dynamic/timbremaps.htm

[35] Pätynen, J. et al.: Anechoic recording system for symphony orchestra. Acta Acustica 94(6), 2008, S. 856–865.