

# Moves & Grooves

Die in den letzten Jahren entwickelten JavaScript-Libraries wie P5, ML5, Plotly sowie verschiedene Bibliotheken zur Audiomerkmalsextraktion ermöglichen die Erstellung völlig neuartiger, robuster und nützlicher Online-Tools zur Datenerhebung. Darüber hinaus sind insbesondere P5/ML5 und Plotly sehr einfach zu erlernen und eignen sich daher vor allem besonders für Studierende, die einen schnellen Einstieg in die Welt der Programmierung suchen. Für den Bereich der musikalischen Akustik sind die Kombinationsmöglichkeiten von synchroner Audiosignalanalyse, Körper-, Hand- und Gesichts-/Ausdruckstracking, sowie von (neuro-)physiologischen Daten besonders wertvoll.

JavaScript als ein bewährtes, nicht-proprietäres und universell im Browser lauffähiges Format bietet dazu eine besondere Nachhaltigkeit, Erweiterbarkeit und Zukunftssicherheit, insbesondere auch durch die Schnittstelle zu Machine-Learning-Modellen (ML5) sowie durch starke und schon langjährig bestehende Communities für P5 und Plotly.

## MODERNE JAVASCRIPT-BIBLIOTHEKEN

P5 ist eine JavaScript-Bibliothek, mit der sich komplexe Programmieraufgaben schnell und einfach lösen lassen.<sup>1</sup> Die Zusatzbibliothek<sup>2</sup> P5.sound ermöglicht die schnelle und unkomplizierte Berechnung von Zeit-Frequenz-Transformationen, Filterungen, Faltungen, Klangsynthese, MIDI und vielen anderen audiobezogenen Verarbeitungsmethoden. Dadurch wird sie für den Bereich der musikalischen Akustik besonders interessant. Mit der 2018 eingeführten Bibliothek ML5 lassen sich innerhalb von P5 vortrainierte Modelle zur Erkennung/Tracking von Personen, Bewegungen, Objek-

ten, Gesichtern, Händen, Tonhöhen und vielem mehr einsetzen.

Nahezu zeitgleich wurde Plotly.js entwickelt<sup>3</sup>. Plotly.js ist eine JavaScript-Bibliothek zur interaktiven Datenvisualisierung, mit der Daten in über 40 verschiedenen Darstellungsarten in 2D und 3D visualisiert werden können. Für die musikalische Akustik ist diese Bibliothek besonders interessant, da in ihr die aus Audiodaten ermittelten Graphen direkt mit den zugehörigen Audiodaten synchronisiert werden können, so dass sich die Ergebnisse nicht nur visuell darstellen lassen, sondern auf interaktive Weise auch hörbar werden.

Mit Meyda<sup>4</sup> und später auch Essentia.js<sup>5</sup>, Aubio.js<sup>6</sup> und Formantanalyzer.js<sup>7</sup> waren spätestens seit 2015 auch umfangreiche Möglichkeiten im Bereich der Audiosignalanalyse in JavaScript verfügbar. So ermöglichen diese Bibliotheken etwa die rechnerische Extraktion einer Vielzahl von Audio-Features wie klanglicher Helligkeit, Rauigkeit, Spectral Centroid, Tonhöhen, Formanten, Lautheit, Vibrato etc. sowie die Transformation des Signals in Chroma- und MFCC-Darstellungen und vieles mehr.

## TOOLS UND ANWENDUNGEN

Sowohl für die Datenvisualisierung als auch für deren Erfassung ermöglichen diese Bibliotheken in ihrer Kombination völlig innovative und vor allem pragmatische, robuste und nachhaltige Lösungen, die im Folgenden beschrieben werden:

### ■ Signalanalyse-Tools

Die Kombination der Bibliotheken P5 und Plotly mit Meyda, Essentia, Aubio oder Formantanalyzer ermöglicht eine schnelle und unkomplizierte Datenerfassung. So wurden mit diesen Bibliotheken interaktive Online-Signalanalyse-Tools er-

stellt, die aus beliebigen hochgeladenen wav- oder mp3-Dateien verschiedenste Audio-Features als interaktive Kurven oder – zur weiteren Verwendung – als JavaScript-Arrays extrahieren können. Die extrahierten Werte lassen sich sowohl als Mittelwerte inkl. Standardabweichung anzeigen als auch als Absolutwerte sowie z-transformiert im CSV-Format zur weiteren Verarbeitung in Excel, JASP, Matlab o. ä. abspeichern. Die Abtastgenauigkeit liegt abhängig von der jeweils verwendeten Signalanalyse-Library zwischen 3 und 50 ±1 ms pro erfassten Wert bzw. von 20 bis 333 Hz.

### ■ Interaktives

#### Valenz-Arousal Modell

Basierend auf der Idee, während des Hörens einer beliebigen hochgeladenen Audiodatei den jeweiligen emotionalen Zustand direkt per Mausbewegung in einem Valenz-Arousal-Feld zu erfassen („EmuJoy“<sup>8</sup>), wurde ein interaktives Valenz-Arousal-Modell entwickelt, das neben den erhobenen Emotionen auch die Audio-Features der gehörten Musikstücke und Klänge oder Geräusche erfasst. Darüber hinaus kann über die Mikrofonbuchse des Computers ein Mindfield@eSense Hautleitwert-Sensor angeschlossen werden, sodass synchron zu den Audio-Features und der emotionalen Selbsteinschätzung auch jeweils ein entsprechender physiologischer Messwert erfasst werden kann. Optional lassen sich auch Videos anstelle der Audiobeispiele abspielen<sup>9</sup> oder die Achsenbezeichnungen für die Erfassung beliebiger anderer Eigenschaften/Einschätzungen ändern. Auch hier können die statistisch vorverarbeiteten Werte (Original, Mittelwerte, SD, z-transformiert) mit einer zeitlichen Auflösung von 50 ±1 ms als CSV-File exportiert werden (s. Abb. 1).

**SINES Tools: Emotion and Audio Features Analysis** (Signal Analysis with [Mevda](#))

(connect a Mindfield edSense SCR sensor via Headset splitter adapter to the microphone input of your computer to record time-synchronous SCR)

Start Upload mp3 Play Data Show JS Arrays Export CSV File

- Click on "Start" to start the experiment.
- Upload an mp3 or wav audio file of your choice
- optional: Select the **noise gate setting** to avoid artifacts in silent parts of the sound:
  - noise gate off (quiet parts will be analysed) or  noise gate on (quiet parts will not be analysed).
- Click on the **center of the valence arousal circle** to start the recording and move your mouse over the fields which correspond most to your mood while listening to music. At the end of the sound example, the path through the valence arousal circle is shown and you can find all data in the data plot below.
- optional: Click on "Play Recording", to follow the recorded path
- Click on "Show JS Arrays" or "Export CSV file", to show the recorded data as Javascript arrays or to export them as a CSV/Excel file

Click on the center of the circle to start the experiment

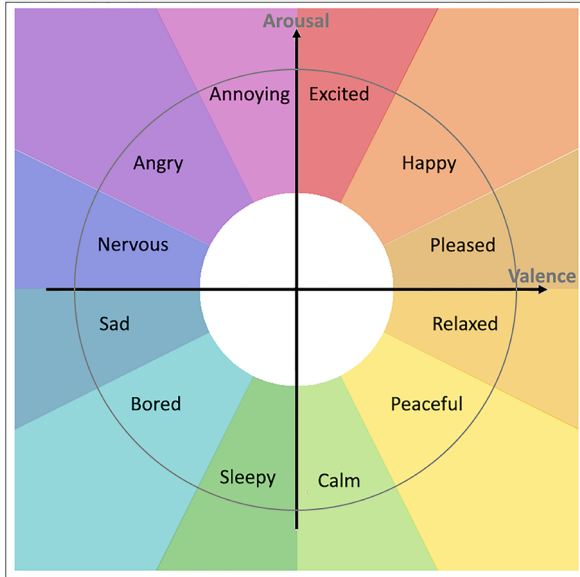


Abbildung 1: Interaktives Valenz-Arousal Feld

**SINES Tools: Video Motion and Audio Feature Analysis** (for one person, works best in Chrome; Signal analysis: [Mevda](#))

Start Upload mp4 Analyse Show JS Arrays Export CSV

- Click on "Start" to set all values on default.
- Upload an mp4 video file showing a moving person.
- optional: Select the **noise gate setting** to avoid artifacts in silent parts of the sound:
  - noise gate off (quiet parts will be analysed) or  noise gate on (quiet parts will not be analysed).
- Click on "Analyse" and play the video to the end.
- optional: **View data** in the plot
- Click on "Show JS Arrays" or "Export CSV" to show the recorded data as Javascript arrays or to export them as a CSV/Excel file.



Abbildung 2: Motion Analysis Tool (Videobeispiel aus <https://youtu.be/vJPM3QoPyNRM> Free HD Video, No Copyright)

**SINES Tools: Facial Expression Analysis** (for one person, works best in Chrome; Signal Analysis: [Mevda](#))

Start Upload mp4 Analyse Show JS Arrays Export CSV

- Click on "Start" to set all values on default.
- Upload an mp4 video file showing the face of a person.
- optional: Select the **noise gate setting** to avoid artifacts in silent parts of the sound:
  - noise gate off (quiet parts will be analysed) or  noise gate on (quiet parts will not be analysed).
- Click on "Analyse" and play the video to the end.
- optional: **View data** in the plot.
- Click on "Show JS Arrays" or "Export CSV" to show the recorded data as Javascript arrays or to export them as a CSV/Excel file.

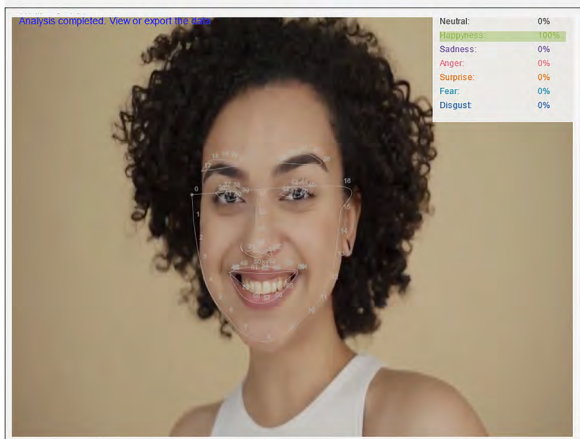


Abbildung 3: Ausdrucks- bzw. Emotionsanalyse (Videobeispiel aus <https://youtu.be/watch?v=QbWStbEv4LU>, Free HD Video, No Copyright)

**SINES Tools: Hand Tracking and Audio Feature Analysis** (for one hand, Chrome only; Signal Analysis: [Mevda](#))

Start Upload mp4 Analyse Show JS Arrays Export CSV

- Click on "Start" to set all values on default.
- Upload an mp4 video file showing a moving hand.
- optional: Select the **noise gate setting** to avoid artifacts in silent parts of the sound:
  - noise gate off (quiet parts will be analysed) or  noise gate on (quiet parts will not be analysed).
- Click on "Analyse" and play the video to the end.
- optional: **View data** in the plot.
- Click on "Show JS Arrays" or "Export CSV" to show the recorded data as Javascript arrays or to export them as a CSV/Excel file.

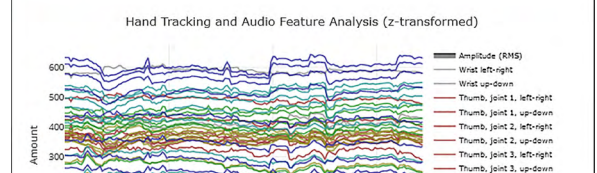


Abbildung 4: Hand Tracking (Videobeispiel aus <https://youtu.be/watch?v=JPu2XQ7X5wI>, Free HD Video, No Copyright)



## ■ Motion Tracking

Über die Machine-Learning-Schnittstelle ML5 lässt sich die Analyse von Audio-merkmalen mit der Erkennung von Bewegungen und Posen von Personen und ihren Körperteilen in Videos kombinieren. Dies eröffnet eine Reihe von Möglichkeiten für die Analyse der Bewegungen von z. B. Tänzer\*innen, Musiker\*innen und Dirigent\*innen außerhalb des Labors (s. Abb. 2). Die X- und Y-Koordinaten aller erfassten Körperteile werden dabei synchron zu den Audio-Features in Zeitabständen von  $100 \pm 2$  ms sowohl als Pixelpositionen als auch z-transformiert erfasst und sind als CSV-Export beliebig weiterverwendbar.

## ■ Facial Expression/Emotion Analysis

Da mit der ML5-Library auch Gesichtsausdrücke erkannt werden können, lässt sich die Audio-Feature-Analyse auch mit einer automatisierten Analyse der Mimik verbinden. Die hierbei erkannten Emotionen bzw. Gesichtsausdrücke werden syn-

chron zur Audioanalyse mit einer Genauigkeit von  $100 \pm 2$  ms aufgezeichnet und können sowohl als Kurven dargestellt als auch im CSV-Format exportiert werden (s. Abb. 3).

## ■ Hand Tracking

Neben den Bewegungen von Personen und der Mimik von Gesichtern können über ML5 auch die Hände von Personen erfasst werden (z. B. von Musiker\*innen beim Spiel ihrer Instrumente). Wie bei den anderen Tools können auch hier die Werte der einzelnen Fingerpositionen synchron zu den entsprechenden Audio-Features als Kurven und/oder im CSV-Format ausgegeben werden (s. Abb. 4).

## ■ Datenlogger für Muse S EEG Headband®-Daten

Mit Hilfe der Web Bluetooth API ([https://developer.mozilla.org/en-US/docs/Web/API/Web\\_Bluetooth\\_API](https://developer.mozilla.org/en-US/docs/Web/API/Web_Bluetooth_API)) kann der Input eines 4-Kanal-EEG-Stirnbandes von Muse® aufgezeichnet werden, um in 4 ms-Intervallen die Bewusstseinszustän-

de von Proband\*innen in ihren Delta-, Theta-, Alpha- und Beta-Bereichen zu bestimmen, während sie Musik oder beliebige andere Audiodateien hören. Zusätzlich werden die Kopfbewegungen aufgezeichnet und die Herzfrequenz gemessen. Aus den gemessenen Daten lassen sich direkt die jeweiligen Oszillationsstärken berechnen und alle Werte können zusammen mit den Audio-Amplituden ihrer klanglichen Umgebung im CSV-Format gespeichert werden.

Die vier EEG-Kanäle und der Herzschlag lassen sich auch sonifizieren, d. h. über einen Synthesizer oder Sampler in Klänge und Geräusche wandeln, so dass man mit Hilfe von diesem Tool und einem Muse S-Stirnband auch mehrstimmig musizieren kann (s. Abb. 5).

Neue Techniken wie Motion Amplification (z. B. zur Blutfluss- und Pulserkennung via WebCam oder zur Erfassung von Korpus-Schwingungen bei Musikinstrumenten in Hochgeschwindigkeits-

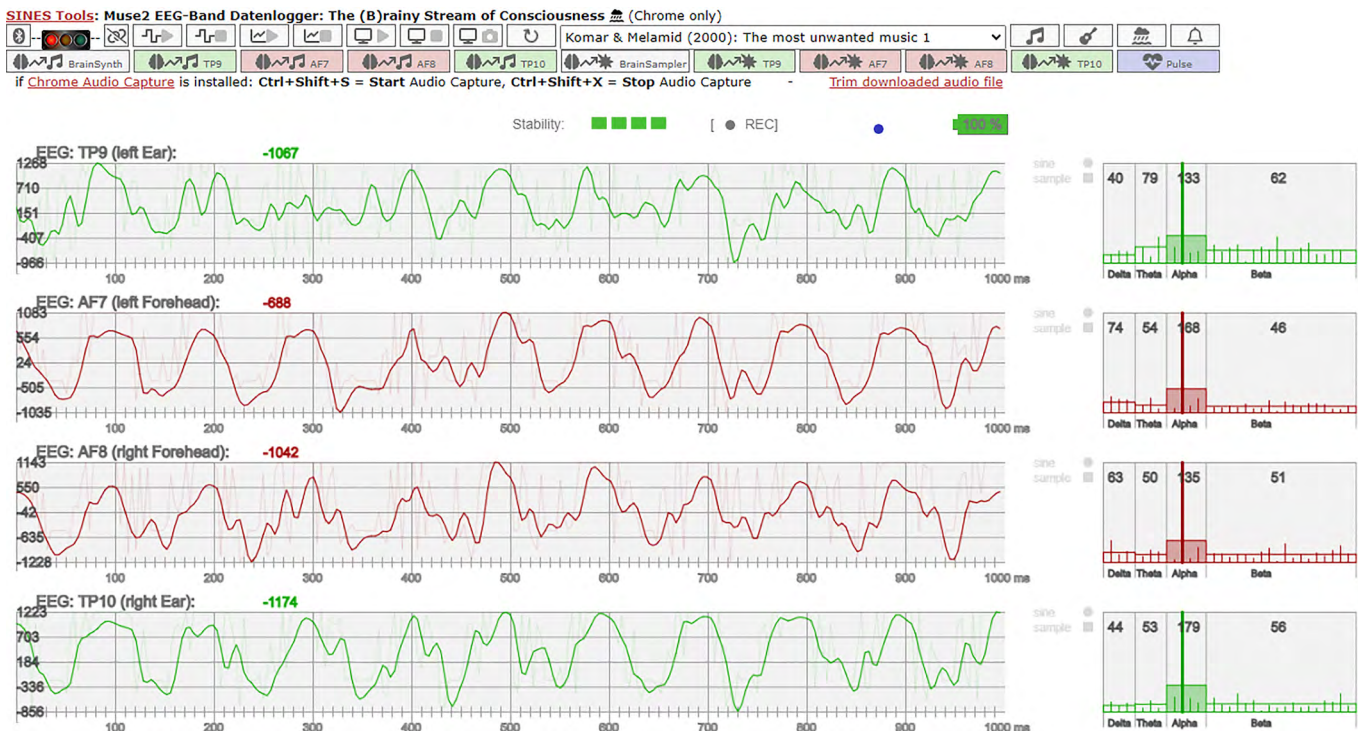


Abbildung 5: EEG-Kurven im Alpha-Zustand, gemessen mit EEG-Datenlogger und Muse S EEG-Stirnband

aufnahmen) sowie die Integration zusätzlicher KI-Modelle (z. B. für genauere Tonhöhen- oder BPM-Einschätzungen oder universelles Pixel-Tracking<sup>10</sup>) werden die Sammlung der hier vorgestellten Werkzeuge in naher Zukunft um weitere

vielfältige und leistungsfähige Anwendungen erweitern.

Alle hier vorgestellten Anwendungen und noch einige weitere Tools finden Sie auf unserer Webseite unter <https://sinestool.univie.ac.at>. Für die Erstellung

des Bluetooth-JavaScript-Handshakes für den EEG-Datenlogger gilt unser besonderer Dank Martin Wittmann.



### Christoph Reuter

ist Professor für Systematische Musikwissenschaft am Musikwissenschaftlichen Institut der

Universität Wien. Seine Forschungsinteressen liegen in der Musikalischen Akustik, Musikpsychologie und Musikinformatik.



### Isabella Czedik-Eysenberg

ist PostDoc-Assistentin am Fachbereich Systematische Musikwissenschaft am

Musikwissenschaftlichen Institut der Universität Wien. Ihre Forschungsinteressen liegen in der Musikpsychologie, Musikinformatik und Audiosignalanalyse.



### Anja-Xiaoxing Cui

ist Tenure Track-Professorin für Neuromedialität am Fachbereich Systematische

Musikwissenschaft am Musikwissenschaftlichen Institut der Universität Wien. Ihre Forschungsinteressen liegen in der Musikpsychologie und -(neuro-)physiologie und der Wechselwirkung zwischen Musik und Sprache.

## Quellen:

1. Reas, C., Fry, B. 2007. Processing: A Programming Handbook for Visual Designers and Artists. Cambridge, Massachusetts: The MIT Press.
2. McCarthy, L., Reas, C., Fry, B. 2015. Getting started with p5.js. Santa Rosa, CA: Make Community, LLC. <https://p5js.org/> <letzter Zugriff: 22.10.2023>
3. Czikk, J. 2013. Ex-Silicon Valley kids to launch web startup in Montreal. Montreal Gazette July 29, 2013. <https://plotly.com/javascript/> <letzter Zugriff: 22.10.2023>.
4. Rawlinson, H., Segal, N., Fiala, J. 2015. Meyda: an audio feature extraction library for the Web Audio API. Proceedings of the first Web Audio Conference (WAC), January 2015, Paris, France. <https://meyda.js.org/> <letzter Zugriff: 22.10.2023>.
5. Correya, A., Marcos-Fernandez, J., Joglear-Ongay, L., Alonso-Jimenez, P., Serra, X. & Bogdanov, D. 2021. Audio and Music Analysis on the Web using Essentia.js. Transactions of the International Society for Music Information Retrieval 4(1), pp. 167-181.
6. Qiuxiang, 2021. Aubios.js, <https://github.com/qiuxiang/aubiojs> <letzter Zugriff: 22.10.2023>.
7. Rehman, A., Liu, Z.T. & Xu, J.-M. 2021. Syllable Level Speech Emotion Recognition Based on Formant Attention. L. Fang et al. (Eds.). Proceedings of the CICA 2021, pp. 261-272. <https://github.com/tabahi/formantalyzer.js> <letzter Zugriff: 22.10.2023>.
8. Nagel, F., Kopiez, R., Crewe, O., Altenmüller, E. 2007. EMuJoy: Software for continuous measurement of perceived emotions in music. Behavior Research Methods 39(2), p. 283-290.
9. Stickler, A., Greil, E., Domnanich, S., Yamada, Y. & Reuter, C. 2023. Entspannend oder verstörend? – Die Wirkung von ASMR mit und ohne Bildmaterial. Fortschritte der Akustik - DAGA2023. 49. Jahrestagung für Akustik 6.-9.3.2023, Hamburg, S. 1220-1222.
10. Wang, Q., Chang, Y.-Y., Cai, R., Li, Z., Hariharan, B., Holynski, A., Snively, N. 2023. Tracking Everything Everywhere All at Once. arXiv:2306.05422

Copyright für die Bilder - alles Screen Copies der beschriebenen Programme - liegt bei Christoph Reuter, Isabella Czedik-Eysenberg und Anja-Xiaoxing Cui; die in den Motion/Gesichts/Handtracking-Tools abgebildeten Bilder von Personen stammen aus Youtube Videos, die extra zur freien Verwendung und ohne Copyright gekennzeichnet sind; die Links sind direkt unter den Bildern.