I see something you can only hear …

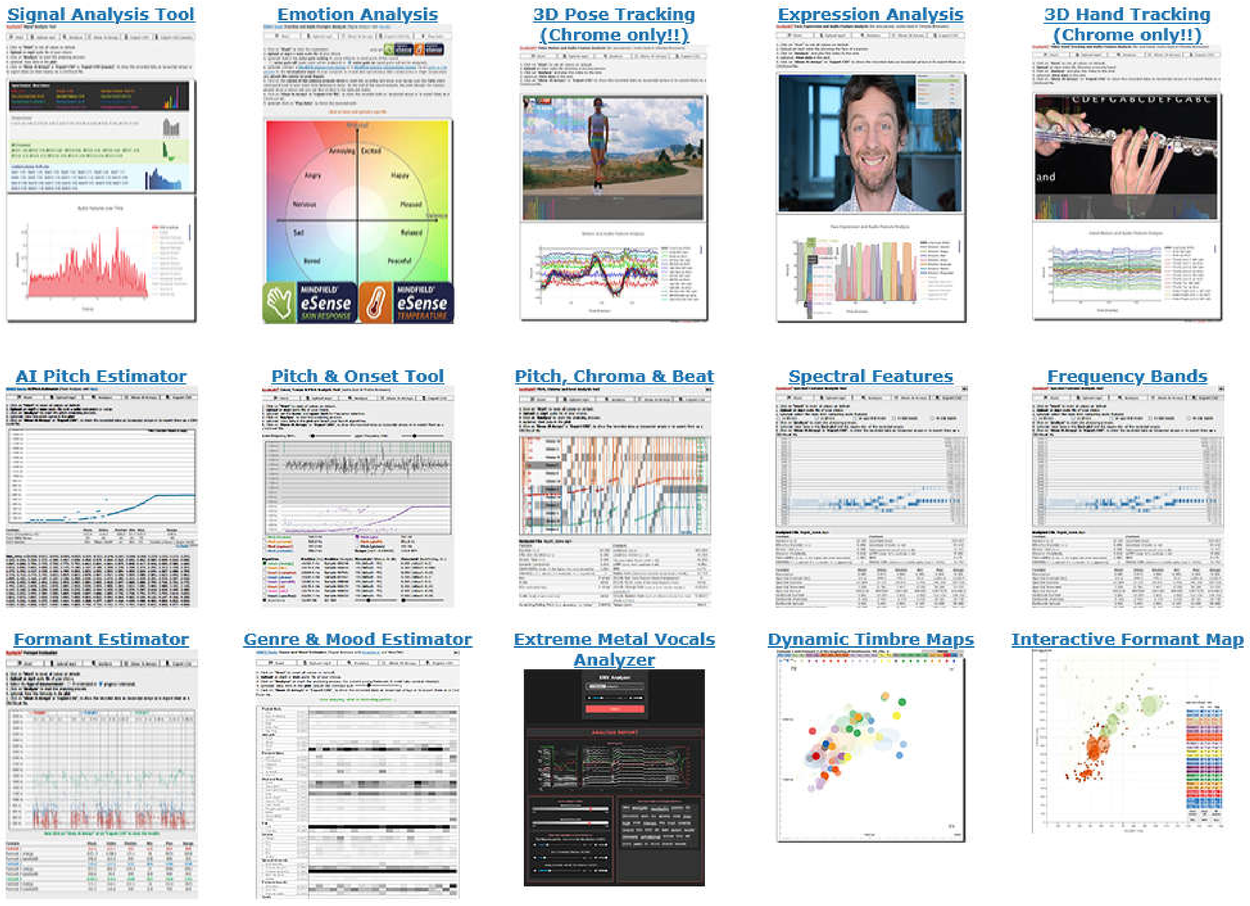
Combining sonic camera recordings with pitch, pose, hand and expression tracking

Christoph Reuter[1], Carolin Schmid[1], Xincan Yang[2], Jian Yang[3]

[1] SInES, Musicological Department, University of Vienna, Austria, [2] Fritz Kreisler Department, University of Music and Performing Arts Vienna, Austria
[3] Department of Music Engineering, Shanghai Conservatory of Music, China

SInES Tools

The current rapid development in the field of audio signal analysis and artificial intelligence as well as the possibilities of the sonic camera expand the possibilities for the visualisation of audio features.
Since 2023, the **S**pace for **In**terdisciplinary **E**xperiments on **S**ound (**SInES**) of the Vienna Systematic Musicology offers a range of free online tools at https://sinestools.univie.ac.at/ with which researchers, students and interested laypersons can gather a variety of music- and movement-related data from audio (wav, mp3) and video (mp4) files (Reuter, Czedik-Eysenberg, Cui 2023a&b).

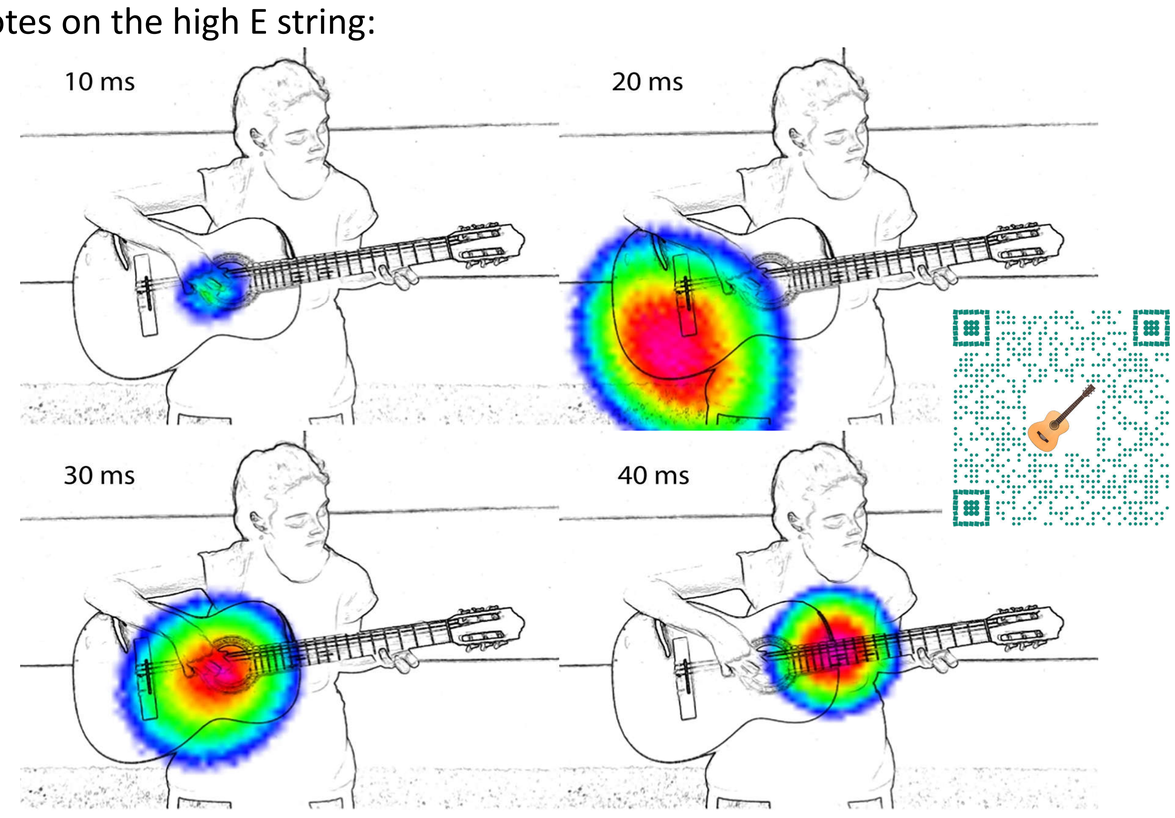


SInES Tools for data acquisition and empirical research (https://sinestools.univie.ac.at)

Combined with videos recorded by a normal or sonic camera, a range of insightful information about musical instruments and their players can be discovered.

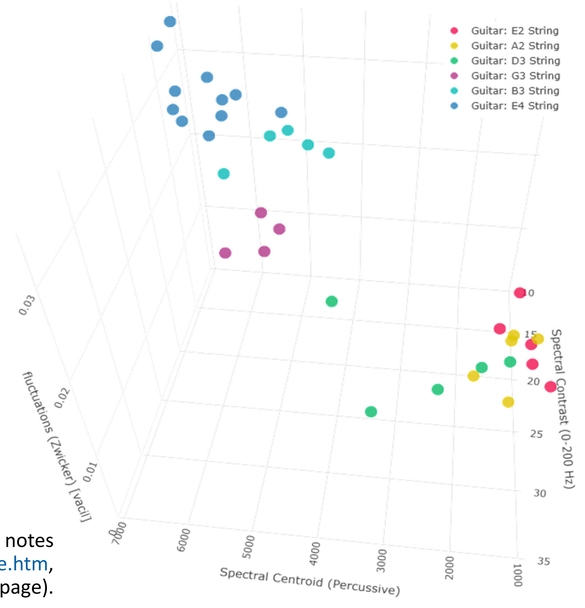
Where on the E string …

With the help of a sonic cam (Gfai Mikado, 96 microphones), the radiation locations for all reachable pitches on a guitar can be visualised and linked to the corresponding audio features via signal analysis, so that the characteristics of the timbres can be described depending on their point of origin. Due to the increasingly shorter wavelengths with rising pitch, this works particularly well for the notes on the high E string:



Sound radiation of the A5 on a guitar's high E-string played by Carolin Schmid at 10, 20, 30 and 40 milliseconds from the beginning of the sound.
The radiation peak travels from the plucked string via the bridge to the body
(all pitches can be played interactively at https://muwiserver.univie.ac.at/sumu2024/gitarre.htm).

Using audio signal analysis, it is also possible to classify the strings of a guitar based on their sound characteristics. The audio features spectral centroid of the percussive part, fluctuations and spectral contrast in the range between 0-200 Hz can be used to classify the strings of the guitar according to their timbral characteristics.



Classification of guitar strings according to the timbral characteristics of their notes (at https://muwiserver.univie.ac.at/sumu2024/gitarre.htm, at the bottom of the page).

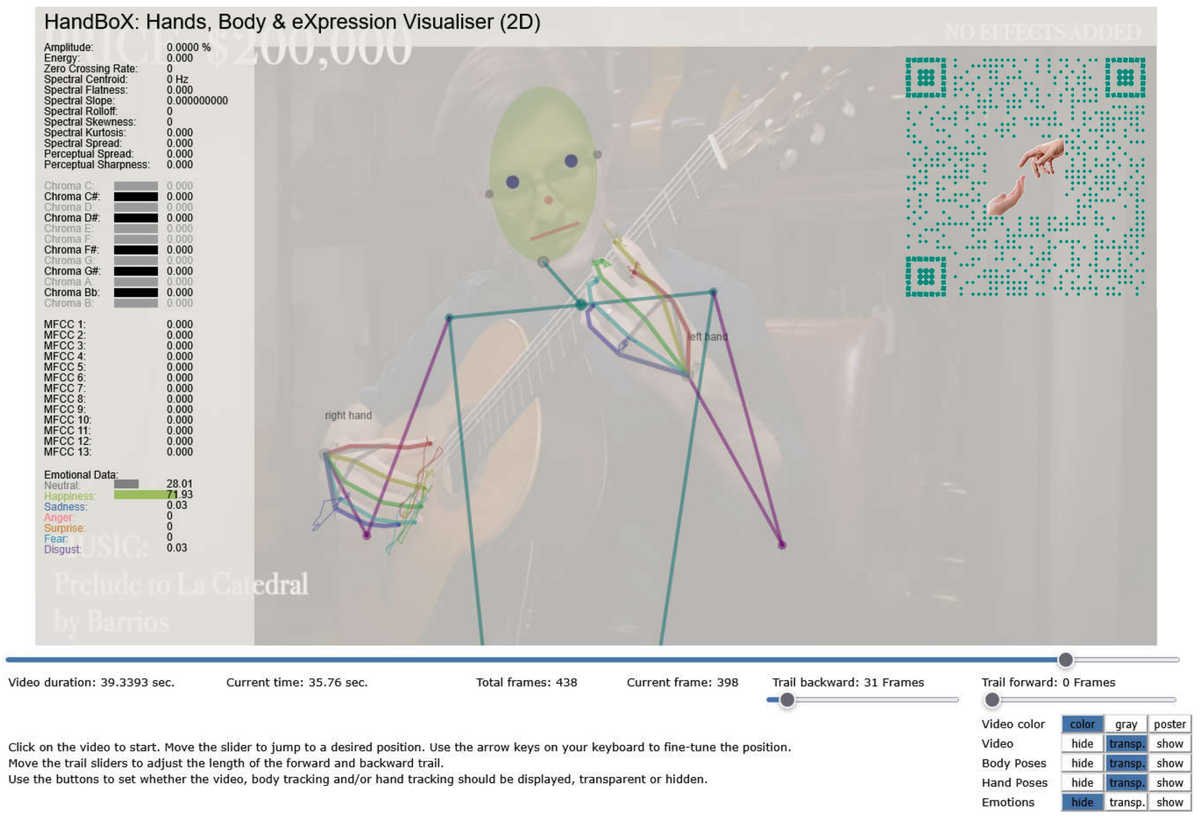
SHAPED and HandBoX

SHAPED (**S**ound Radiation, **H**and, **A**natomy, **P**itch & **E**xpression **D**isplay)
Merging the data obtained with the SInES tools for body, pitch, hand and expression tracking based on sonic camera and normal camera recordings. Thanks to the pitch tracking model used in SInES Tools (Google CREPE, Kim et al. 2018), even fast-played notes can be determined with a high degree of accuracy. Arrow keys can be used to jump from note to note in order to intuitively visualise information about the body and hand position, sound radiation, intonation, even players expression (shown by color in the face) as well as a range of other timbral parameters for each segment of the piece.

SHAPED: Sound Radiation, Hand, Anatomy, Pitch & Expression Display: Combination of data from SInES tools and acoustic camera with display options in the menu at the bottom right as well as additional displays for amplitudes, expression/emotion of the player and recognised selectable pitches underneath. (Violin played by Silas Yang at https://muwiserver.univie.ac.at/sumu2024/violin_data.htm)

HandBoX (Hand Body eXpression Visualizer)

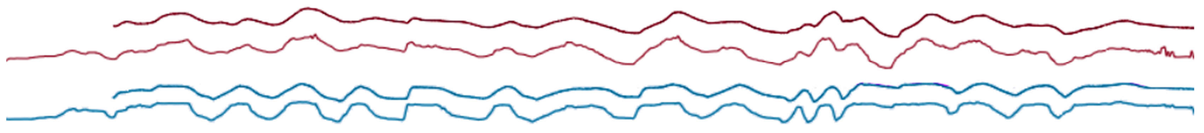
The HandBoX visualiser has a similar structure and can be used to display a combination of body and hand movement, players emotion and audio data collected with the SInES tools.



HandBoX (Hand Body eXpression Visualizer): displaying body, hand, expression and audio data of musical instruments and their players (Brandon Acker from https://www.youtube.com/watch?v=EPY9Mxg3OJQ, in https://muwiserver.univie.ac.at/sumu2024/handbox.htm).

Summary

The data collected with the SInES tools can be visualised in many different ways, with or without a sonic camera. The combination of the sonic camera and the SInES tools opens up new possibilities for finding connections between sound radiation, audio features and musician movements.



Movement curves of a violinist's right hand (played by Jian Yang at https://muwiserver.univie.ac.at/sumu2024/violin_MoCap.htm):
top (**red**): left-right movement: Qualisys MoCap system (100 fps) and SInES tool (pose-tracking Samsung mobile phone video, 10 fps)
bottom (**blue**): up-down movement: Qualisys MoCap system (100 fps) and SInES tool (pose-tracking Samsung mobile phone video, 10 fps)

Literature

Kim, J.W., Salamon, J., Li, P., Bello, J.P., (2018). CREPE: A Convolutional Representation For Pitch Estimation. *arXiv*:1802.06182. <> Reuter, C., Czedik-Eysenberg, I., Cui, A.-X. (2023a). Happy Life comes with P5 - P5, ML5, Meyda and Plotly as helpful Tools in Teaching and Research. *Proceedings of DAGA2023, 49. German Annual Conference on Acoustics* (p. 991-994). Hamburg. <> Reuter, C., Czedik-Eysenberg, I., Cui, A.-X. (2023b). Moves & Grooves. Analyse von Audiosignalen, Tracking von Körper, Hand und Gesicht/Ausdruck.. *OCG Journal* 48/4, S. 20-23