Science rules with SInES Tools - New Updates, New Apps

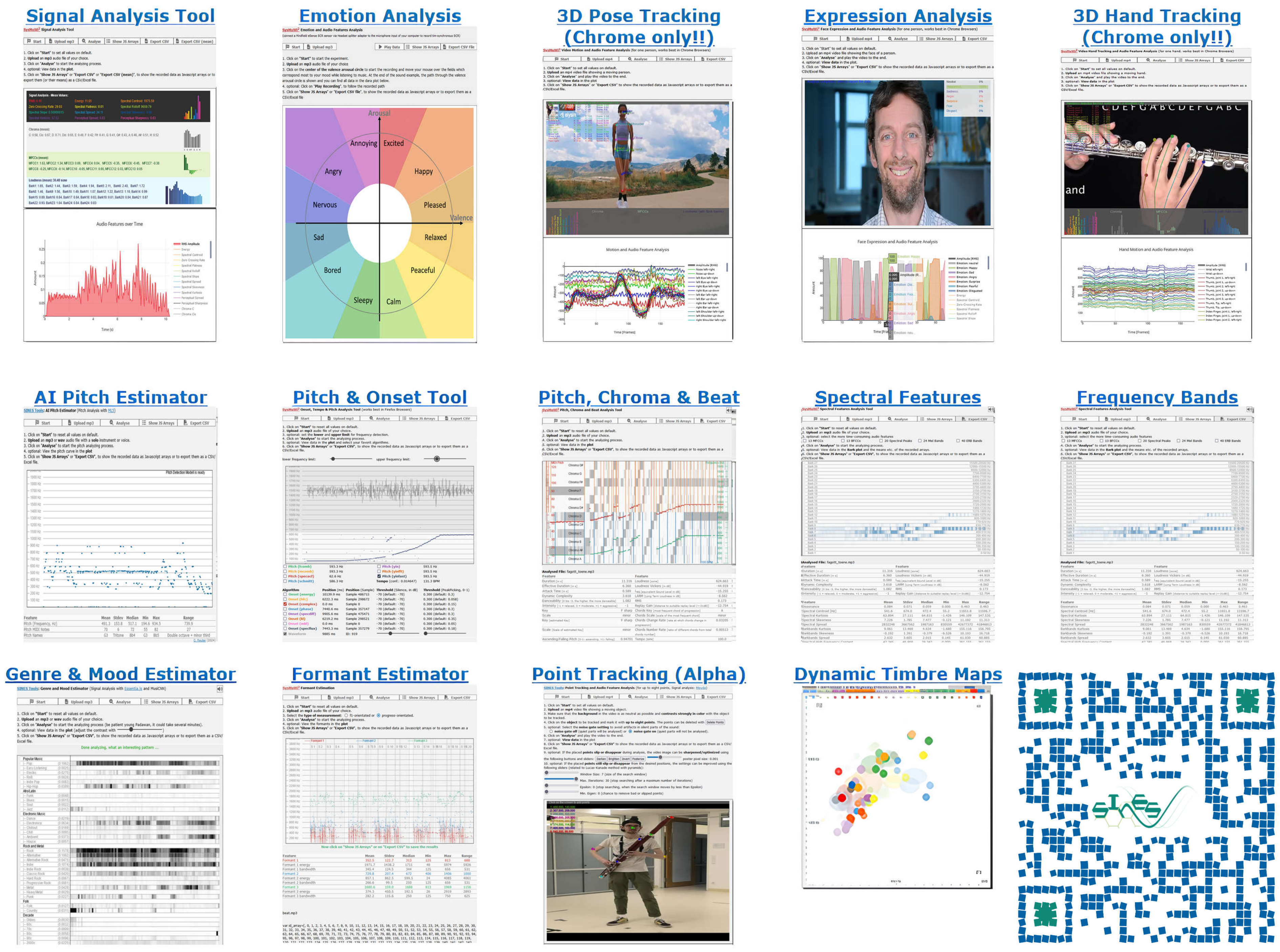
Christoph Reuter[1], Isabella Czedik-Eysenberg[1], Anja-Xioaxing Cui[1], Marik Roos[1], Matthias Eder[1], Jian Yang[3], Matthias Bertsch[2]

[1] SInES, Musicological Department, University of Vienna, Austria
[2] Motion-Emotion Lab, University of Music and Performing Arts, Vienna, Austria
[3] Department of Music Engineering, Shanghai Conservatory of Music, China

40. Annual Conference of the German Society for Music Psychology (DGM)
Munich, 7.-9. September 2024

SInES Tools

Since 2023, the Space for Interdisciplinary Experiments on Sound (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)

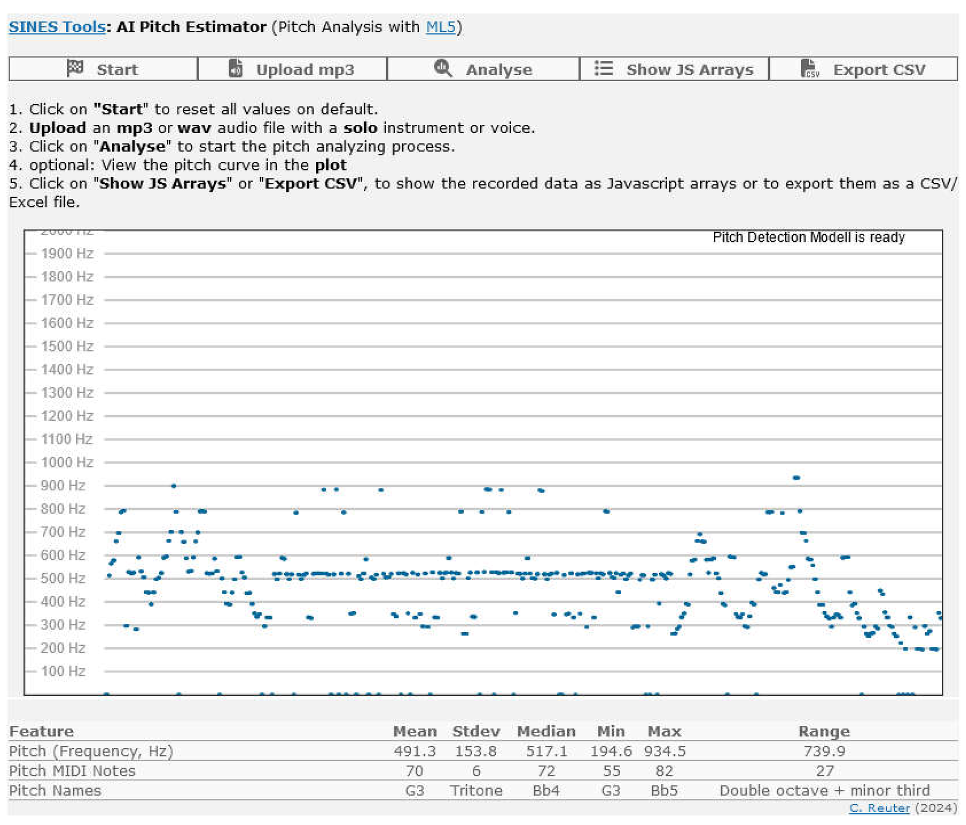
New Updates, New Apps

In recent months, a number of updates have been implemented that have made the tools even more user-friendly, stable and productive:

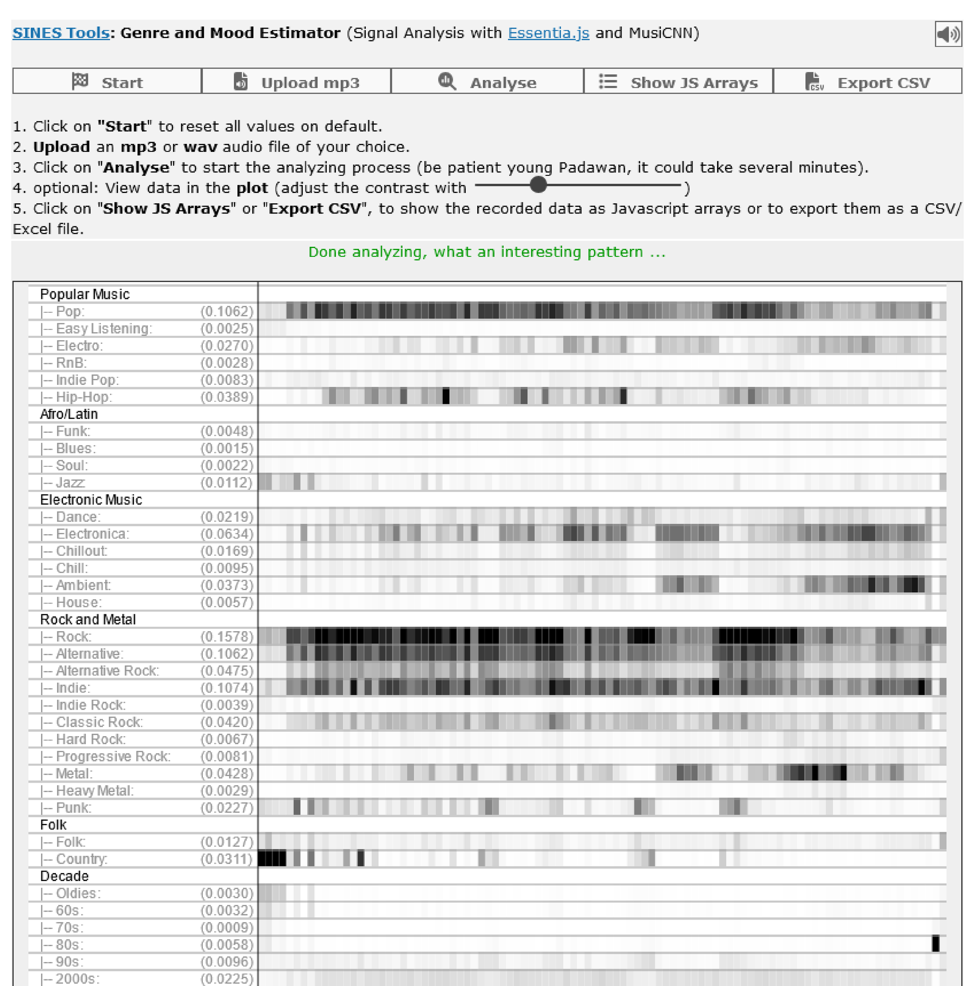
- Pose/Hand/Expression Tracking: Integration of new AI models that now allow more precise detection of movements in three dimensions.
- Pose/Hand/Expression Tracking: Adaptation of the detection rate to the frame rate (fps) of the uploaded video (fewer drop-outs).
- Hand Tracking: Robustness in the assignment of handedness (no more two left hands ;->)
- Emotion Analysis: device-independent connection of the Mindfield eSense SCR sensor using an audio-to-USB adapter.

Based on pre-trained TensorFlow models for pitch recognition (CREPE, Kim et al. 2018) and for genre and mood estimation (Alonso-Jiménez et al. 2020), pitch and genre & mood estimator apps were developed as new SInES Tools:

The **Pitch Estimation** app can capture monophonic melody lines below 2000 Hz more accurately than the pitch tracking algorithms currently in use. The detection takes place in real time with a resolution of < 20 ms. In addition to the frequencies, it also displays the corresponding musical pitches and MIDI numbers.



SInES Tools: AI Pitch Estimator (top) and Genre & Mood Estimator (bottom)

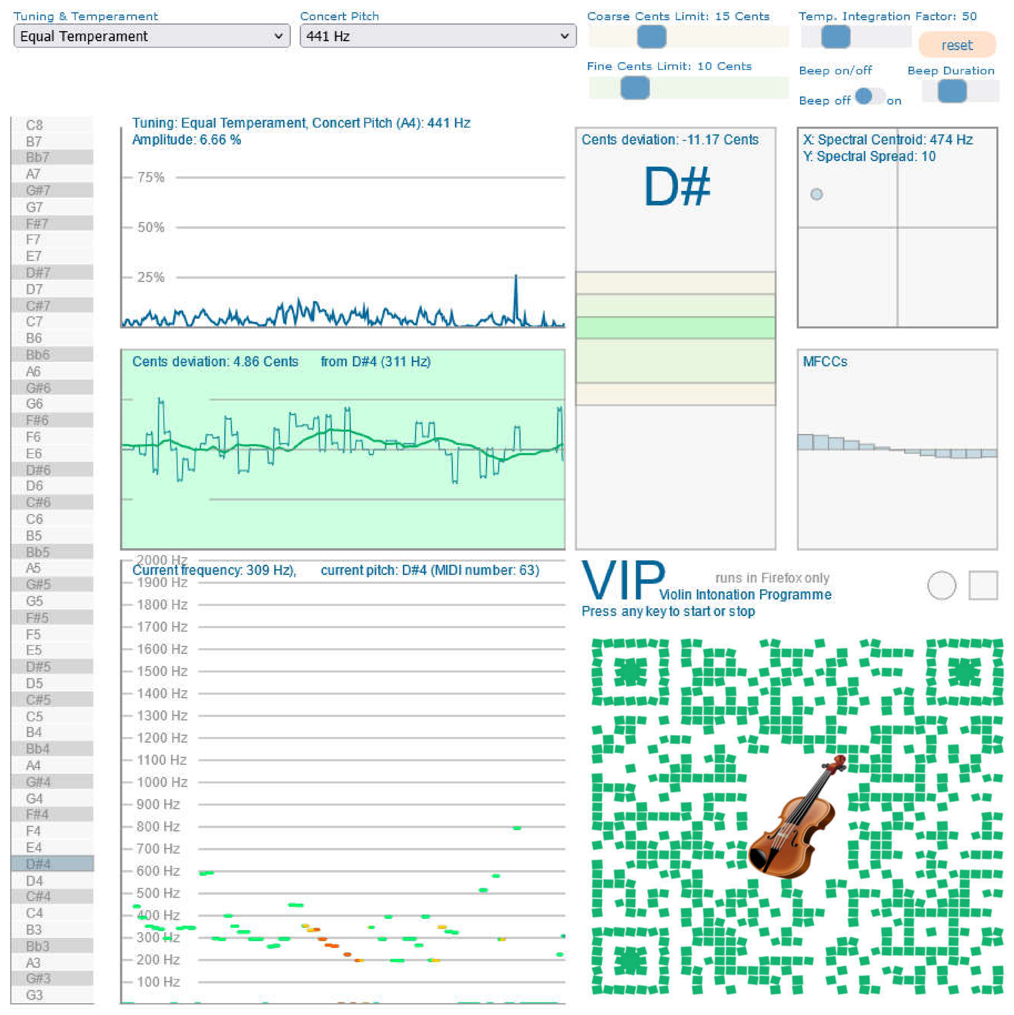


The **Genre & Mood Estimator** rates the extent to which a piece of music corresponds to a certain genre from the rock/pop/jazz area, how danceable/catchy/sexy/beautiful it is, who is singing (man, woman or instrumental), what mood the piece conveys (happy/sad/relaxed/aggressive) etc. in time windows of 3 seconds each.

In both apps (like in all SInES Tools), the calculated data is z-transformed, analysed for mean value, median, standard deviation etc. and can be exported as Javascript arrays and CSV files.

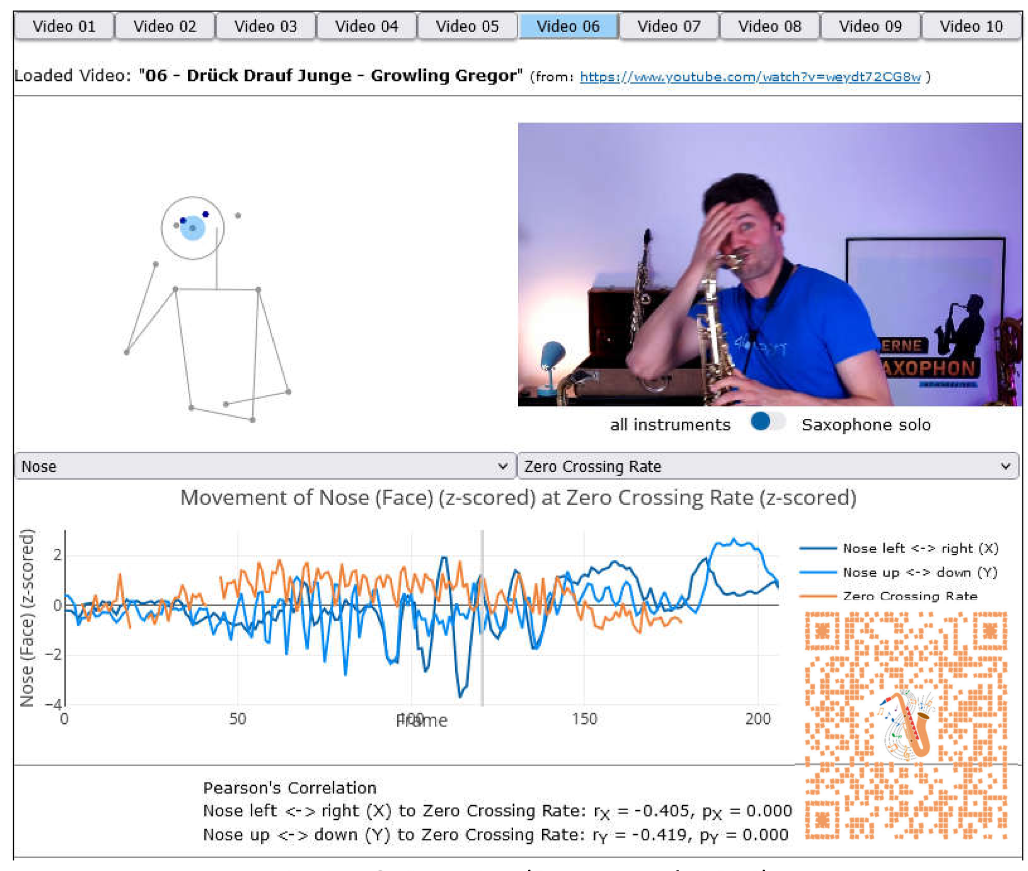
...even more Apps

VIP (Violin Intonation Program): Application for intonation control while playing the violin, including adjustment of the concert pitch and the tuning system as well as error tolerance, vibrato smoothing, score calculation for intonation accuracy, etc.



VIP Violin Intonation Program (Reuter, Kölsch 2024)
https://sinestools.univie.ac.at/vip.htm

Moving & Grooving: Application that can be used to search for correlations between musician movements and audio features. The application draws line figures of the movements recorded via pose tracking and lists all significant correlations found.



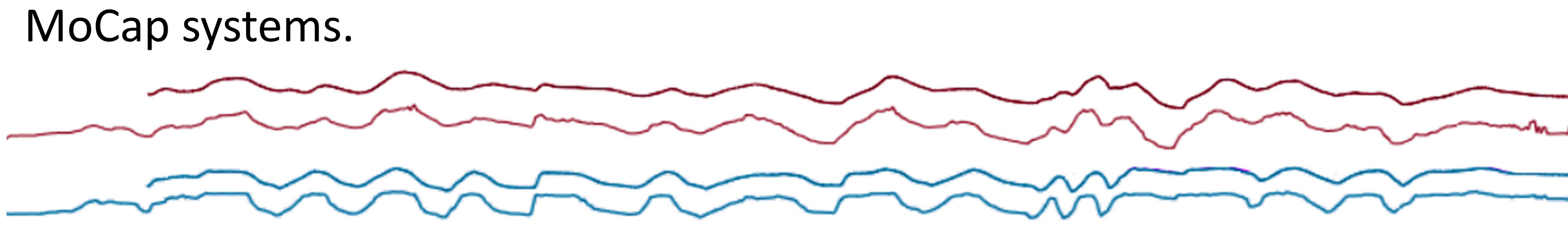
Moving & Grooving (Reuter et al. 2024)
https://muwiserver.univie.ac.at/moving_grooving

SHAPED (Sound Radiation, Hand, Anatomy, Pitch & Expression Display): 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 here, even fast-played notes can be determined with a high degree of accuracy and the 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, expression and a range of other tonal parameters for each segment of the piece.

SHAPED (Sound Radiation, Hand, Anatomy, Pitch & Expression Display)
https://muwiserver.univie.ac.at/sumu2024/violin_data.htm

Summary

Thanks to new AI models, existing SInES Tools have been improved and new tools have been developed. The SInES Tools are now even more frame-accurate, more robust and also more user-friendly than the first versions. Pose and hand tracking tools can now track movements in three dimensions and the interactive Valence/Arousal model now also allows the skin conductance sensor to be connected to the PC using a corresponding audio-to-USB adapter, regardless of the equipment.
In addition to Pitch and Genre & Mood Estimator tools, specialized applications such as the Violin Intonation Programme (VIP) or the Sound Radiation, Hand, Anatomy, Pitch & Expression Display (SHAPED) could also be created, whereby the recorded movement curves (left/right, up/down) are comparable to the corresponding movement curves of professional MoCap systems.



Movement curves of a violinist's right hand:
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.
- Alonso-Jiménez, P., Bogdanov, D., Pons, J., Serra, X. (2020). Tensorflow Audio Models in Essentia. ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP).
- 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.