Representing expressive performance in tempo-loudness space

Jörg Langner[1)] & Werner Goebl[2)]

[1)] Musikwissenschaftliches Seminar,
Humboldt-Universität zu Berlin
Unter den Linden 6
D-10099 Berlin, Germany
Tel: +49 30 2093 2065
Fax: +49 30 2093 2183
Email: jllangner@aol.com

[2)] Austrian Research Institute for
Artificial Intelligence (ÖFAI)
Schottengasse 3
A-1010 Vienna, Austria
Tel: +43 1 5336112 24
Fax: +43 1 5336112 77
Email: werner.goebl@ai.univie.ac.at

Background

A large number of very detailed studies analyse various parameters of expressive music performance (e.g., timing, dynamics, articulation. For an overview see Palmer, 1997; Gabrielsson, 1999). They often fall short by addressing each expressive parameter separately. In our every day experience, we perceive these parameters simultaneously and the impression of a performance always results from an integrated perception of all performance parameters. Given that these parameters generally interact with each other, an integrated analysis technique of more than one parameter at a time may provide new insight into an individual performer's expressive strategies, and into the phenomenon of musical expression generally.

Another problem of performance analysis is the enormously large amounts of information the researcher has to deal with, even when investigating e.g. only the timing of a few bars for a single piece. It remains generally unclear whether the expressive deviations measured are due to deliberate expressive strategies, motor noise or imprecision of the performer, or even measurement errors. So smoothing over a longer time span is an appropriate means to reduce local variability in data. In this way, higher-level patterns of expression could be visualised and immediately evaluated for music performance research.

Aims

In the present contribution, we develop an integrated analysis technique in which tempo and loudness are processed and displayed simultaneously. Both the tempo and loudness curves are smoothed with large window size (corresponding to the length of a bar). These two performance parameters are then displayed in a two-dimensional performance space on a computer screen: a dot moves in synchrony with the sound of the performance. The trajectory of its tail describes geometric shapes that may be intrinsically different for different performers. In this way, we get something like the performer's "finger print" for a given realisation of a piece. A set of expressive performances was collected to better understand and evaluate the meaning of such emerging patterns, they were compared with aesthetic ratings of these performances by musically trained listeners.

Method

The timing information was taken from annotated MIDI data of expressive piano performances played on a Bösendorfer computer-controlled grand piano (SE290). The loudness information is derived from the corresponding audio files (represented in the perceptual measure of loudness [sone], see Zwicker & Fastl, 2001; Langner, 2002). Both tempo and loudness deviations were smoothed over a large window (using overlapping Gaussian windows e.g. at the bar level, corresponding to 3–4 seconds, see Langner, Kopiez, Stoffel, & Wilz, 2000). Performance animations were created that display these data in synchrony with the music: a dot moves through a two-dimensional space of tempo (*x* axis) against loudness (*y* axis), leaving behind it a trajectory (for a snapshot see Figure 1).

The investigated data were 22 expressive performances of two excerpts by Chopin (from Op. 10/3 and Op. 38, data described in detail in Goebl, 2001) played by professional pianists, and, for each excerpt, one average performance (Goebl, 1999). Some of these performances have been aesthetically rated by 72 musically trained listeners in an extensive listening test (the test will not further be described in this context). For each performance, its expression trajectory was analysed in the context of its aesthetic rating in order to better understand their properties.

To compare these performances to a famous concert pianist, a recording by Maurizio Pollini[1] was chosen. The tone onsets for this demonstration example were measured by hand from the audio file, from that the procedure was as above.

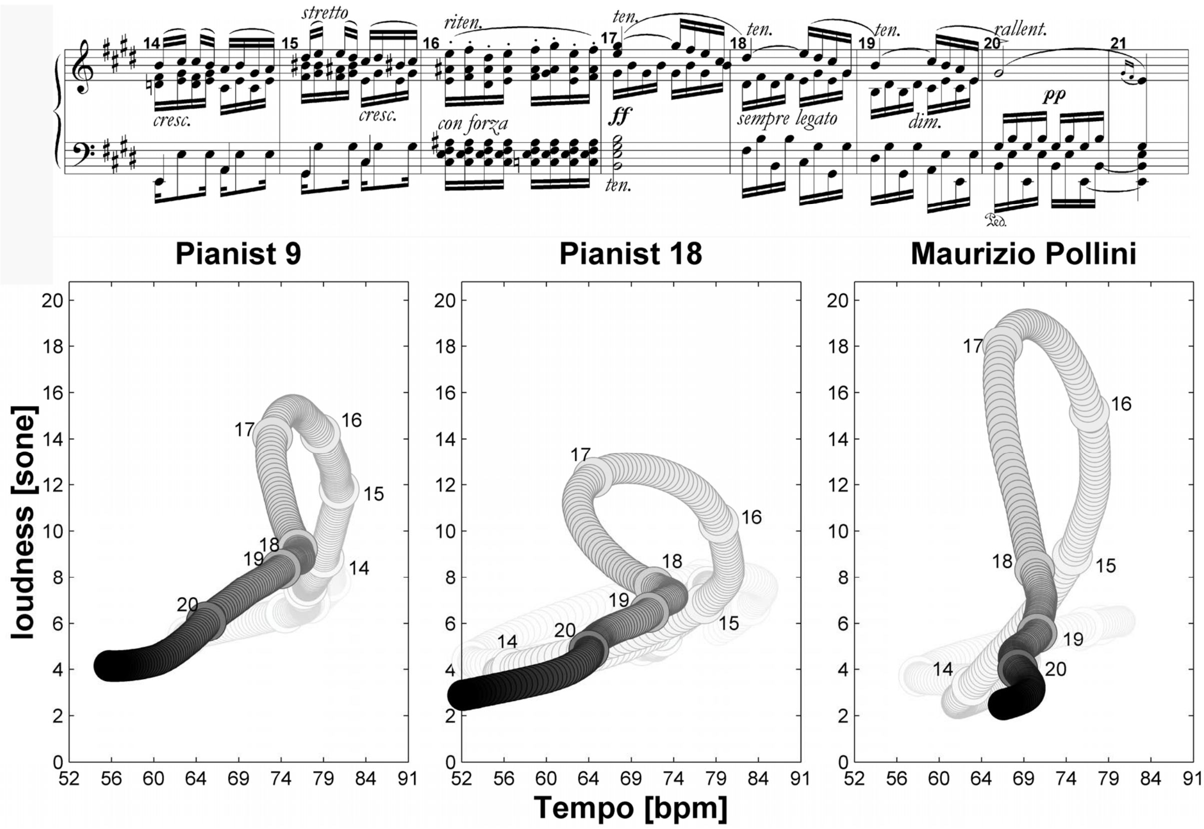

Figure 1. Expression trajectories over the last bars (14-21) of the beginning section of Chopin's etude op. 10, no. 3 performed by Pianist 9 (left), Pianist 18 (middle), and Maurizio Pollini (right). *x* axes: tempo in beats per minute, *y* axes: loudness measured in sone. The darkest points represent the end of the excerpt, while instants further in the past appear fainter. The more prominent circles indicate the beginning of a new bar and its number.

[1] Maurizio Pollini, Chopin Etudes, Deutsche Grammophon, 413 794-2, recorded 1972.

Results

The Figure shows part of the trajectories of three different performances with the relevant part of the musical score (top), the bar numbers correspond to the numbers in the graphs. The performance of Pianist 9 (left) was rated relatively low, in contrast to Pianist 18 (middle) whose performance was rated very high. Pollini's example was not included in the rating experiment (the recording would have sounded very different to the homogeneous recordings of the Chopin excerpts), but it is shown here to exemplify different expressive strategies.

The most fundamental similarity between all performances is that the expressive trajectory tends to go to the lower left side of the space at phrase boundaries (which corresponds to a slowing down and a decrescendo towards the end of phrases). This strategy is used more extensively by Pianist 18 than by Pianist 9. Another striking commonality between the pianists is that at the beginning of a phrase the performers increase first the tempo and then the intensity.

In the excerpt shown here (measures 14-21), all three pianists show a anticlockwise loop that reflects the strong *crescendo* and the hold back (*ritenuto*) of the climax, as indicated in the score. They all did not play loudest at the point marked *fortissimo* (bar 17), but got softer before the notated fortissimo. However, the shapes of the loops displayed in Figure 1 reveal certain differences: while Pianist 9 already starts at a relatively loud and fast stage (bar 14), both Pianist 18 (the well-rated performance) and Pollini started their crescendo in the lower left corner of the performance space (slow and soft). Especially these measures were pointed out by the listeners as being particularly well played by pianist 18. The artistry of Pollini comes out when the concept of his interpretation is taken into consideration: for the first 14 bars he remains very soft and in a quite steady tempo to reserve his expressive energies for the outburst in the displayed section.

Another difference between the famous pianist and the two others is that Pollini does not slow down too much at the end of the section. He planned, of course, to play of whole Etude and not only the first 21 measures, as the other pianists did.

Discussion

A novel approach to visualising two performance parameters in a two-dimensional space was introduced shortly in this paper. In the spoken presentation also the animation will be shown, where the expressive trajectory builds up over time parallel to the music. The usefulness of this technique was tested with performances of two short excerpts by Chopin. It remains to be investigated what kinds of trajectories we will get with other types of music, and whether the observations made with the Chopin performances could be generalised.

We envisage a range of future applications of this technique. Apart from further research on musical expression, a more complex computational implementation of this technique may be used to automatically analyse existing audio recordings. It could also be used as an on-line tool in music teaching, analysis, and practice.

Acknowledgements

This research was in part supported by the project Y99-INF, sponsored by the Austrian Federal Ministry of Education, Science, and Culture (BMBWK) in the form of a START Research Prize. The Austrian Research Institute for Artificial Intelligence acknowledges basic financial support from the Austrian Federal Ministry for Education, Science, and Culture.

References

Gabrielsson, A. (1999). Music Performance. In D. Deutsch (Ed.), *Psychology of Music, second edition* (pp. 501-602). San Diego: Acadamic Press.

Goebl, W. (1999). Analysis of piano performance: towards a common performance standard?, *Proceedings of the Society of Music Perception and Cognition Conference (SMPC99)*. Evanston, Illinois, USA.

Goebl, W. (2001). Melody lead in piano performance: Expressive device or artifact? *Journal of the Acoustical Society of America, 110*(1), 563-572.

Langner, J. (2002). *Musikalischer Rhythmus und Oszillation. Eine theoretische und emprirische Erkundung. Including a comprehensive abstract in English*. Frankfurt a. M.: P. Lang.

Langner, J., Kopiez, R., Stoffel, C., & Wilz, M. (2000). Realtime analysis of dynamic shaping. In C. Woods & G. Luck & R. Brochard & F. Seddon & J. A. Sloboda (Eds.), *Proceedings of the 6th international conference on music perception and cognition*. Keele, UK: Keele University Department of Psychology.

Palmer, C. (1997). Music performance. *Annual Review of Psychology, 48*, 115-138.

Zwicker, E., & Fastl, H. (2001). *Psychoacoustics. Facts and Models*. Berlin, Heidelberg: Springer.