What Really Moves Us in Music: Expressivity as a Challenge to Semantic Audio Research

Gerhard Widmer[*]

Department of Computational Perception
Johannes Kepler University Linz, Austria
gerhard.widmer@jku.at

Abstract

Semantic Audio and Music Information Retrieval (MIR) Research have led
to some spectacular successes in recent years. They have made it possi-
ble to build machines that listen to music, follow and track it in real time,
extract useful information like melody, harmonies, beat and rhythm, cate-
gorise music into genres and other classes, automatically monitor TV and
radio stations worldwide, and do many other useful things. All of this is
based on algorithms that analyse audio and extract information from the
audio signal that relates to what we may call the 'semantics' of the signal:
higher-level aspects that most humans readily perceive when listening to
musical audio.

Despite these successes, however, I would claim that we have barely
scratched the surface of what music really *is*, what it means, and what it
is in music that moves us. In this presentation, I would like to focus on
a dimension of the semantics of musical audio that has not yet received
much attention in the semantic audio research community: the dimension
of expressivity in musical performance – especially in the performance of
classical music. Expressivity in performance refers the millions of different
ways in which skilled performers can play the same piece of music, shaping

---

it and making it come alive via subtle nuances in tempo, timing, dynamics, articulation, timbre, etc., and in this way giving the piece a unique character and emotional quality and expressing their own understanding and personal 'meaning' of the music. That this is an important aspect of the music is evidenced by the fact that millions of music lovers all over the world continue to buy recordings of (or go to see yet another live performance of) pieces they have already heard many times.

In this presentation I will

- introduce performance expression as an important dimension of the 'semantics' of a piece of music;

- discuss and demonstrate what it means to extract, quantify, characterise, and model this via machine;

- review recent work on computational modelling of performance expression;

- sketch possible application scenarios;

- and identify pertinent research challenges for our field.

Generally, my aim is to promote a field of research that not all attendees of this conference may be so familiar with and that, I believe, is of real importance when we speak of the 'semantics' of musical audio.