

## Music Performance Panel: NICI / MMM Position Statement

Peter Desain, Henkjan Honing and Renee Timmers

*Music, Mind, Machine* Group

NICI, University of Nijmegen

mmm@nici.kun.nl, www.nici.kun.nl/mmm

In this paper we will put forward our view on the computational modeling of music cognition with respect to the issues addressed in the *Music Performance Panel* held during the MOSART 2001 workshop. We will focus on issues that can be considered crucial in the development of our understanding of human performance and perception in its application to computer music systems. Furthermore, they were chosen such as to complement the issues brought forward by the other contributing institutes (i.e. OFAI/Vienna, KTH/Stockholm, and DEI/Padua). In summary these are:

- A computational model in agreement with music performance data is *starting point* of research, rather than an *end product* (cognitive modeling is preferred over a descriptive model)
- Importance of empirical data obtained in controlled experiments (rather than using individual examples of music performances)
- Preference for the concept of *performance space* (over the use of large corpora of music performances)
- Study performance through perception, focusing on the constraints of expression rather than studying the ideal or “correct” performance (as such avoiding the issue of performance style, and enabling the study of important aspects that are not directly measurable in the performance data itself, e.g., those of a perceptual and/or cognitive nature)

## **Research aims**

The panel addresses a number of dichotomies in the study of music performance, such as theory-driven vs. data-driven, oriented towards cognitive plausibility vs. computational simplicity, perception-oriented vs. production-oriented. The discussion aims to reveal research aims and methods, which are quite varied among research groups.

In our group, we study music perception and performance using an interdisciplinary approach that builds on musicology, psychology and computer science (hence the name *Music, Mind, Machine*). The aim is to better understand music cognition as a whole. The method is to start with hypotheses from music theory, to formalize them in the form of an algorithm, to validate the predictions with experiments, and, often, to adapt the model (and theory) accordingly. In other words, in the method of *computational modeling*, theories are first formalized in such a way that they can be implemented as computer programs. As a result of this process, more insight is gained into the nature of the theory, and theoretical predictions are, in principle, much easier to develop and assess. With regard to computational modeling of musical knowledge, the theoretical constructs and operations used by musicologists are subjected to such formalization. Conversely, with computational modeling of music cognition, the aim is to describe the mental processes that take place when perceiving or producing music, which does not necessarily lead to the same kind of models. As such, for us, a computational model that mimics human behavior is not enough. It in fact is more a starting point of analysis and research, than an end product (see [1] for an elaborate description).

## **Evaluation and validation of music performance models**

One of the key issues in developing algorithms and computational models is their validation on empirical data. In the case of the MOSART project, music that is artificially generated should respect human perception and performance such as to assure seamless interaction and intelligible control by its users. For evaluating and validating models of expression, it is problematic to search for a “correct”, general or

benchmark interpretation of music [2], to which the models can be compared. Though this approach is quite common in AI modeling, it is very unattractive for music cognition research. Not only is the notion of an ideal performance questionable, comparing the input-output relation between the model and the musical performance is also too limited an evaluation. A data-driven perspective might eventually result in an accurate description [2,3], it will, however, not be a model, in the cognitive sense. It needs to describe more than just an input-output transformation. In fact, a good model is a model for which changes in parameter settings that relate to manipulated aspects of the performance (e.g. by instruction to the performer) remains to show agreement between model and performance. As such step by step further validating the model.

As an illustration of the difference between a model and a good description from another domain, one can take difference between FM-synthesis and physical modeling. It is possible to generate very convincing sounds with FM synthesis (after careful selection of the parameters). However, the whole space of sounds is unintuitive and difficult to control. In contrast, physical models have more similarity with the human world and succeed in replicating the behavior of existing objects (e.g., made of tubes and strings) that are known to the user and are therefore easier to control, despite their more restricted expressive power.

In general, a computational model that captures important aspects of human perception and action will be more successful in computer music systems. Models that simply aim at an input-output agreement do not necessarily give us a better understanding of the underlying perceptual or cognitive processes, which is essential for the development of convincing and intuitive models for human interaction with machines (see [4] for a discussion on the psychological validation of models of music cognition). A solely data-driven approach ignores the fact that important aspects of music performance are not directly measurable or present in the data itself. For instance, tempo (or expressive rubato for that matter) is a percept, and cannot be directly measured. The same applies for syncopation and other temporal aspects of music that exist due to (violations of) listener's expectations.

With regard to the methodology of evaluating models of expression, we assign great importance to the systematic collection of empirical data, experimentally

manipulating the relevant parameters. For instance, in our research on expressive vibrato [5, 6], we explicitly control for global tempo to reveal how it is adapted to the duration of notes. And we record repeated performance to get a better grip on consistency (e.g. to be able to separate between intended and non-intended expressive information).

Similarly, in our studies on piano performances (e.g., [7]), only careful experimental manipulation of a few parameters (like global tempo, or the addition or removal of one note) will give a precise insight in the underlying mechanisms that we need to reveal in order to make better computer music editing software or music generation systems.

Blindly examining very large samples of music performance is clearly not an alternative to this.

And, finally, in our work in rhythm perception, we put quite some effort in developing methods that allow us to investigate the concept of *performance space*, abstracting from individual examples. The idea here is to consider all possible interpretations, including musical and unmusical ones, in a variety of styles. While currently we only applied this approach to relatively short fragments of music [8], we find this method a more systematic and insightful alternative for randomly grown corpora of music performances. In addition, studying the perception of rhythm is also a way to identify the constraints on expressive timing in music performance (instead of focusing on an ideal or unique performance) as such avoiding the notion of a “correct” performance, which is an important advantage that allows for models to be elaborated independent of performance style.

## References

- [1] Documents on <http://www.nici.kun.nl/mmm> under heading “Research”.
- [2] Sundberg, J., Friberg, A., and Frydén, L. (1991) Common Secrets of Musicians and Listeners: An Analysis-by-Synthesis Study of Musical Performance. In P. Howell, R. West & I. Cross (eds.). *Representing Musical Structure*. London: Academic Press.
- [3] Widmer, G. (2001) Using AI and Machine Learning to Study Expressive Music Performance: Project Survey and First Report. *AI Communications*, 14.

- [4] Desain, P., Honing, H., Van Thienen, H. & Windsor, L.W. (1998). Computational Modeling of Music Cognition: Problem or Solution? *Music Perception*, 16 (1), 151-16.
- [5] Desain, P. & Honing, H. (1996) Modeling Continuous Aspects of Music Performance: Vibrato and Portamento [ICMPC Keynote address], *Proceedings of the International Music Perception and Cognition Conference*. CD-ROM, Montreal: McGill University.
- [6] Rossignol, S., Desain, P. & Honing, H. (2001). State-of-the-art in fundamental frequency tracking. *Proceedings of the Workshop on Current Research Directions in Computer Music*. Barcelona: UPF.
- [7] Timmers, R., Ashley, R., Desain, P., Honing, H., and Windsor, L. (in press) Timing of ornaments in the theme of Beethoven's Paisiello Variations: Empirical Data and a Model. *Music Perception*.
- [8] Desain, P. & Honing, H. (submitted). *The Perception of Time: The Formation of Rhythmic Categories and Metric Priming*. See <http://www.nici.kun.nl/mmm/time.html>