Tracking expressive performances with linear and non-linear timing models

Sensori-motor synchronization (SMS) describes the ability of humans to rhythmically coordinate movements to external stimuli. SMS can be viewed from two theoretical perspectives: According to the dynamical systems theory, SMS involves non-linear phase and period adjustments to a set of coupled internal oscillators. The information-processing approach posits linear phase and period correction of the internal timekeeper. Models derived from these theories are commonly tested in non-musical tapping experiments. In music performance tempo fluctuations are used to increase expressivity which makes tracking of musical sequences more difficult for these models. In order to improve the tracking capabilities, we propose and test an extension to both models by introducing piece-specific tempo expectations. This extension is tested on a corpus of expressively performed music. A set of symbolic performance data containing excerpts of Chopin's piano etude Op. 10 No. 3 performed by 22 professional pianists comprises the test corpus. Tempo expectations are modeled using local inter-onset intervals (averaged across performances). We test four different models: 2 (model types, linear/non-linear) x 2 (information on tempo expectations, y/n). Two methods are used: first, we optimize each model to fit individual performances so that the lowest possible timing errors (between predicted and actual note onset times) are achieved. Second, we run a cross-validation experiment to test the generalization capabilities of the models. For this purpose each model is optimized to fit a subset (60%) of the performances and the evaluation is done on the remaining performances. This procedure is repeated multiple times with random assignments to the optimization/evaluation subsets. The experiments are currently under way. Preliminary results suggest that adding information on typical expressive timing patterns improves model predictions. We plan to further test these models in a musical synchronization task in which a human co-performs with a virtual duet partner driven by the extended models.