A Computational Study of the Role of Tonal Tension in Expressive Piano Performance

Carlos Eduardo Cancino-Chacón[1,2] and Maarten Grachten[2]

[1]Austrian Research Institute for Artificial Intelligence, Vienna, Austria
[2]Institute of Computational Perception, Johannes Kepler University Linz, Austria

JOHANNES KEPLER UNIVERSITY LINZ

OFAI

Using a computational approach, we study how three measures of tonal tension contribute to the prediction of expressive performances of classical piano music. We use non-linear sequential models (recurrent neural networks) to assess the role of these measures in predicting expressive tempo and dynamics for a dataset of Mozart piano sonatas performed by a professional concert pianist (Roland Batik).

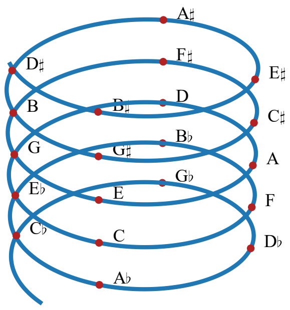
Tonal Tension Features

Chew's (2000) spiral array model is a 3D representation of pitch classes, chords, and keys, for which spatial proximity indicates tonal relationships.

Cloud diameter (T_{cd}): maximal tonal distance between notes

Cloud momentum (T_{cm}): harmonic movement as the tonal distance from one section to the next

Tensile strain (T_{ts}): relative tonal distance between the current segment and the center of effect of the key



Herremans & Chew (2016) showed that these features correlate to human perception of tonal tension, by comparison to Farbood (2012).

Score Features

Pitch (P)

(a) *pitch$_h$, pitch$_l$, pitch$_m$*: chromatic pitch of the highest, lowest, and melody note at each score position

(b) *vic$_1$, vic$_2$, vic$_2$*: vertical intervals above the lowest note in a chord

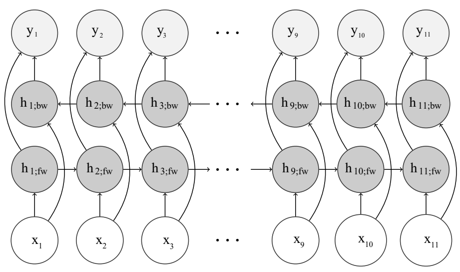
Metrical (M)

(a) $b_{\Phi,t}$: relative location of an onset in a bar

(b) b_d, b_s, b_w: metrical strength at the downbeat, secondary strong beat (in duple meter), and weak metrical positions

Model

Our model architecture consists of a composite bidirectional long short-term memory layer (LSTM) with multiplicative integration and a linear dense layer with one unit as output.



Experiments

Eight 5-fold cross validation experiments were run for each expressive parameter, using models trained on all combinations of features.

Tempo

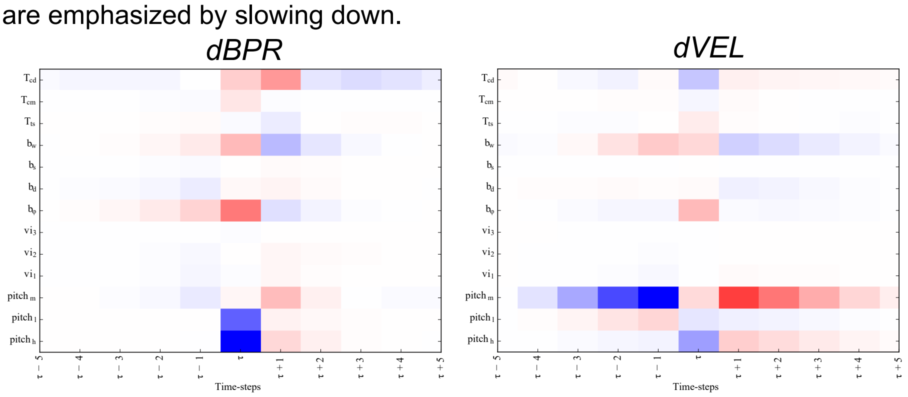
BPR and *dBPR*: local beat period ratio and its 1st derivative (indicating change in tempo)

Dynamics

VEL and *dVEL*: max. performed MIDI velocity per score position and its 1st derivative (indicating change in loudness)

Results & Discussion

Experiments of models trained with and without tonal tension features show that tonal tension helps to predict *change* in tempo and dynamics more so than absolute tempo and dynamics. We ran a sensitivity analysis, which shows that "high-tension" passages in performance are emphasized by slowing down.



Differential sensitivity maps. Red indicates that the feature contributes positively (e.g., it increases MIDI velocity, increasing loudness, or makes beat period longer, thereby reducing tempo). Blue indicates that the feature contributes negatively. Each cell indicates feature contributions per time step away from the currently performed note.

Table values are average R^2 over the corresponding 5-fold cross-validation. Green indicates that tension features improve predictions; red indicates that tension features worsen predictions. Bold values are statistically significant.

Tempo

Feature Set	BPR		dBPR		
		+ T		+ T	d
Ø	-	0.010	-	0.011	
P	0.024	0.021	0.068	**0.073**	(0.10)
M	0.051	0.054	0.093	0.092	
P + M	0.056	0.054	0.105	**0.110**	(0.06)

Dynamics

Feature Set	VEL		dVEL		
		+ T		+ T	d
Ø	-	0.018	-	0.026	
P	0.326	0.335	0.236	**0.250**	(0.16)
M	0.048	0.052	0.041	**0.050**	(0.18)
P + M	0.351	0.347	0.250	**0.282**	(0.40)

Acknowledgements: Funding for this work was provided by the European Research Council under the EU's Horizon 2020 Framework Programme (ERC Grant Agreement number 670035, project "Con Espressione".