CON ESPRESSIONE!
AI, Machine Learning, and Musical Expressivity

Gerhard Widmer
Institute of Computational Perception, Johannes Kepler University Linz (JKU)
AI Lab, Linz Institute of Technology (LIT)
ELLIS Unit Linz

JOHANNES KEPLER
UNIVERSITY LINZ

European Research Council
Established by the European Commission

AI & MUSIC?

A LITTLE EXPERIMENT

JKU

IDENTIFY THE HUMAN PIANIST

Friedrich Kuhlau, *Allegro Burlesco*, Op.88 No.30

(1) (2) (3)

JKU

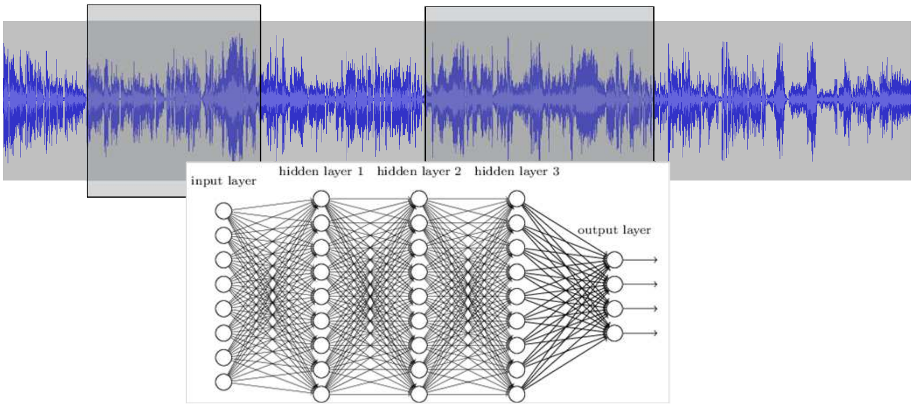
AI & MUSIC

JKU

COMPUTATIONAL MUSIC PERCEPTION

JKU

COMPUTATIONAL MUSIC PERCEPTION

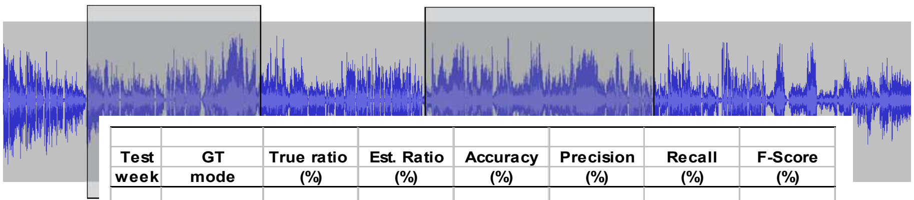


MUSIC DETECTION

JKU

* Schlüter, J. (2017). *Deep Learning for Event Detection, Sequence Labelling, and Similarity Estimation in Music Signals*. PhD Thesis, Johannes Kepler University Linz, Austria, 2017.

COMPUTATIONAL MUSIC PERCEPTION



Test week	GT mode	True ratio (%)	Est. Ratio (%)	Accuracy (%)	Precision (%)	Recall (%)	F-Score (%)
1	min	77,63	77,39	96,66	98,00	97,69	97,84
1	max	78,95	77,39	96,59	98,80	96,86	97,82
1	ave	78,29	77,39	96,62	98,40	97,27	97,83
1	same	78,67	77,79	97,25	98,79	97,69	98,24
2	min	80,09	79,43	96,73	98,36	97,54	97,95
2	max	82,07	79,43	96,21	99,27	96,09	97,65
2	ave	81,08	79,43	96,47	98,81	96,81	97,80
2	same	81,70	80,29	97,40	99,27	97,54	98,40
3	min	74,42	73,89	94,13	96,38	95,70	96,04
3	max	77,04	73,89	94,30	98,27	94,26	96,22
3	ave	75,73	73,89	94,21	97,33	94,97	96,13
3	same	76,43	74,45	95,40	98,24	95,70	96,95

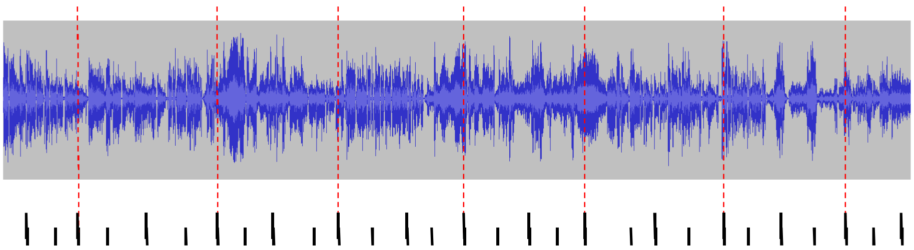
SWISS PERFORM

MUSIC I

JKU

* Schlüter, J. (2017). *Deep Learning for Event Detection, Sequence Labelling, and Similarity Estimation in Music Signals*. PhD Thesis, Johannes Kepler University Linz, Austria, 2017.

COMPUTATIONAL MUSIC PERCEPTION



RHYTHMIC STRUCTURE: Beat, Tempo, Measures/Bars, Rhythm
MUSIC DETECTION

* Krebs, F., Holzapfel, A., Cemgil, A.T. and Widmer, G. (2016). Inferring Metrical Structure in Music Using Particle Filters. *IEEE Transactions on Audio, Speech & Language* 23(5), 817-827.

COMPUTATIONAL MUSIC PERCEPTION

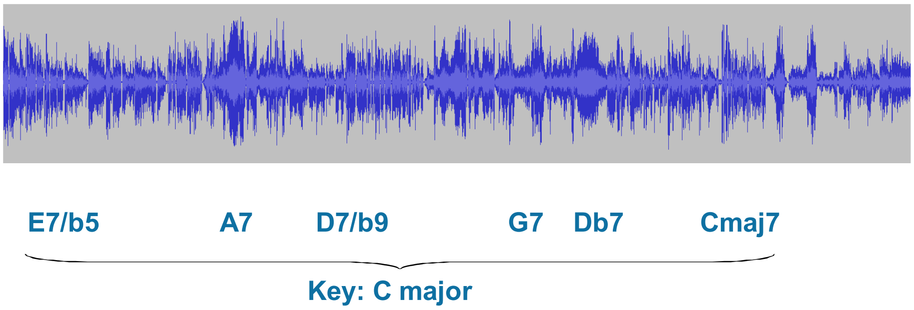
Florian Krebs

RHYTHMIC STRUCTURE: Beat, Tempo, Measures/Bars, Rhythm
MUSIC DETECTION

IEEE Signal Processing Cup 2017
ICASSP 2017, New Orleans

* Krebs, F., Holzapfel, A., Cemgil, A.T. and Widmer, G. (2016). Inferring Metrical Structure in Music Using Particle Filters. *IEEE Transactions on Audio, Speech & Language* 23(5), 817-827.

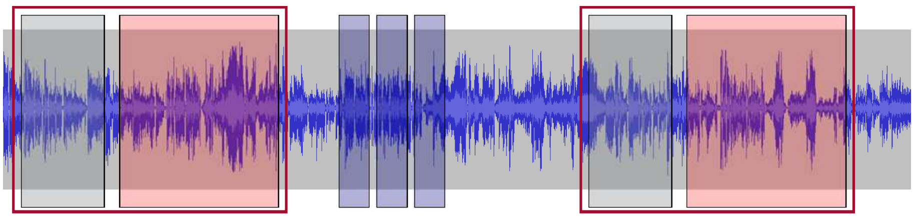
COMPUTATIONAL MUSIC PERCEPTION



HARMONIC STRUCTURE: Key, Chords, Chord Sequences
RHYTHMIC STRUCTURE: Beat, Tempo, Measures/Bars, Rhythm
MUSIC DETECTION

* Korzeniowki, F. (2018).
Harmonic Analysis of Musical Audio using Deep Neural Networks.
PhD Thesis, Inst. of Computational Perception, Johannes Kepler University Linz (JKU).

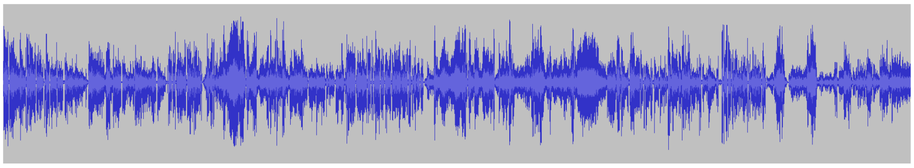
COMPUTATIONAL MUSIC PERCEPTION



SEGMENT STRUCTURE: Chorus, Verse, Song Boundaries, …
HARMONIC STRUCTURE: Key, Chords, Chord Sequences
RHYTHMIC STRUCTURE: Beat, Tempo, Measures/Bars, Rhythm
MUSIC DETECTION

* Grill, T. and Schlüter, J. (2015).
Music Boundary Detection Using Neural Networks on Combined Features and Self-similarity Lag Matrices. In *Proceedings of EUSIPCO 2015*, Nice.

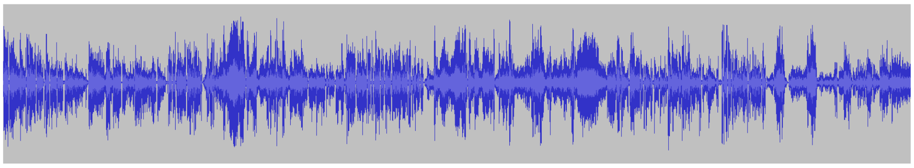
COMPUTATIONAL MUSIC PERCEPTION



GENRE, STYLE: Jazz / Rock / Pop / Folk / HipHop / ...
SEGMENT STRUCTURE: Chorus, Verse, Song Boundaries, ...
HARMONIC STRUCTURE: Key, Chords, Chord Sequences
RHYTHMIC STRUCTURE: Beat, Tempo, Measures/Bars, Rhythm
MUSIC DETECTION

JKU

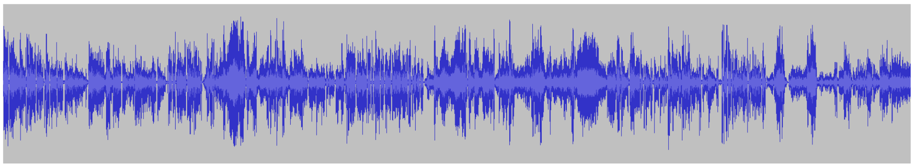
COMPUTATIONAL MUSIC PERCEPTION



GENERAL MUSICAL SIMILARITY (e.g., for recommendation)
GENRE, STYLE: Jazz / Rock / Pop / Folk / HipHop / ...
SEGMENT STRUCTURE: Chorus, Verse, Song Boundaries, ...
HARMONIC STRUCTURE: Key, Chords, Chord Sequences
RHYTHMIC STRUCTURE: Beat, Tempo, Measures/Bars, Rhythm
MUSIC DETECTION

JKU

REAL-TIME MUSIC TRACKING

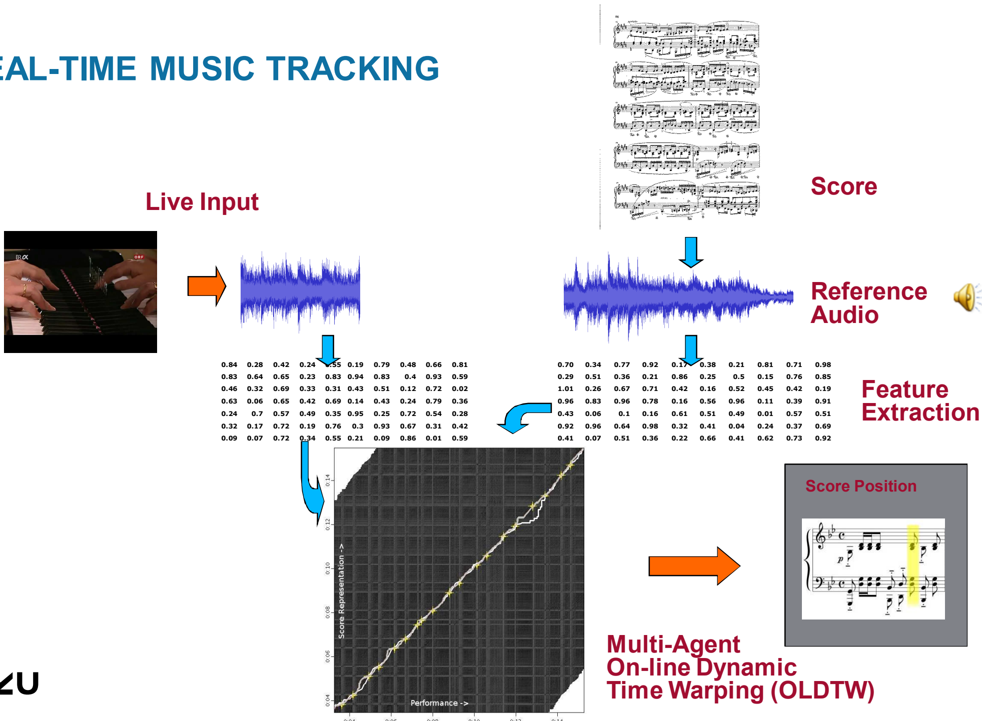


Video

JKU

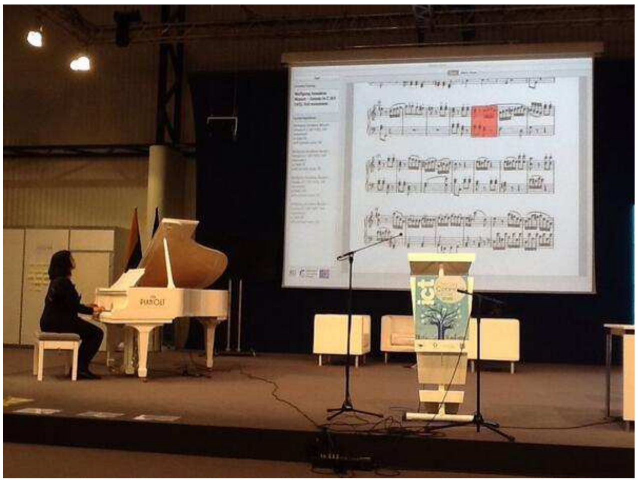
* Arzt, A. (2017). *Flexible and Robust Music Tracking.*
PhD Thesis. Johannes Kepler University, Linz, Austria, 2016.

REAL-TIME MUSIC TRACKING



JKU

REAL-TIME PIECE IDENTIFICATION AND TRACKING

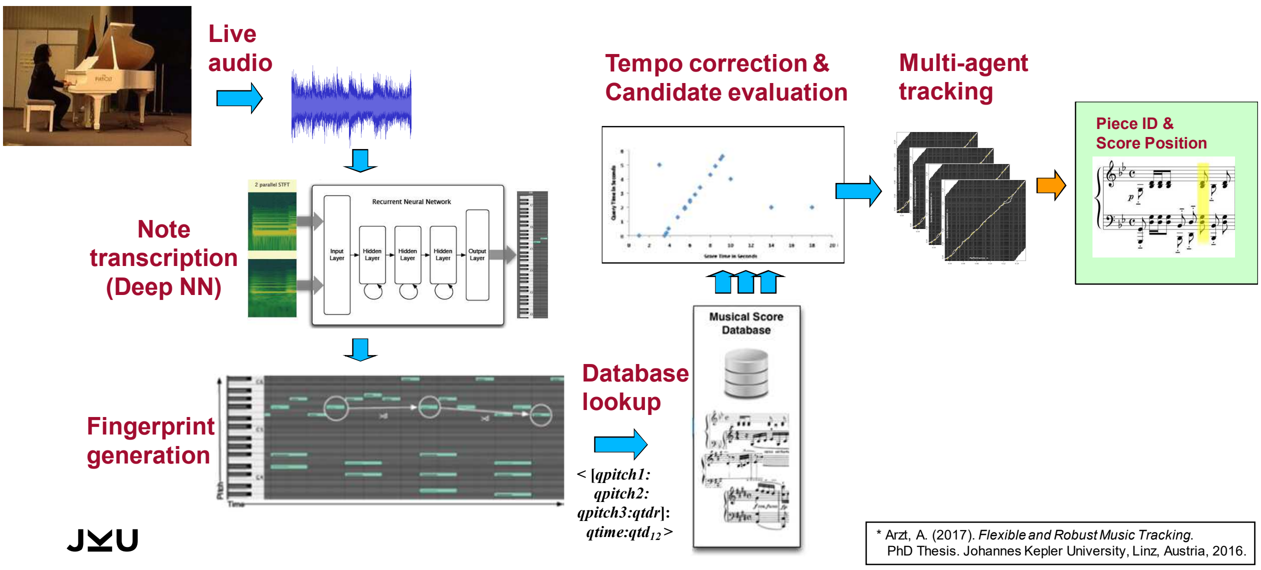


Pianist:
Cynthia Liem

JKU

* Arzt, A. (2017). *Flexible and Robust Music Tracking.* PhD Thesis. Johannes Kepler University, Linz, Austria, 2016.

REAL-TIME PIECE IDENTIFICATION



JKU

* Arzt, A. (2017). *Flexible and Robust Music Tracking.* PhD Thesis. Johannes Kepler University, Linz, Austria, 2016.

CONCERTGEBOUW AMSTERDAM, Dec. 20, 2014

Royal Concertgebouw
Orchestra, Amsterdam

Mariss Jansons, conductor

* Arzt, A., Frostel, H., Gasser, M., Grachten, M. & Widmer, G. (2015). Artificial Intelligence in the Concertgebouw. In *Proceedings of the 24th International Joint Conference on Artificial Intelligence (IJCAI 2015)*, Buenos Aires, Argentina.

JKU

CONCERTGEBOUW AMSTERDAM, Dec. 20, 2014

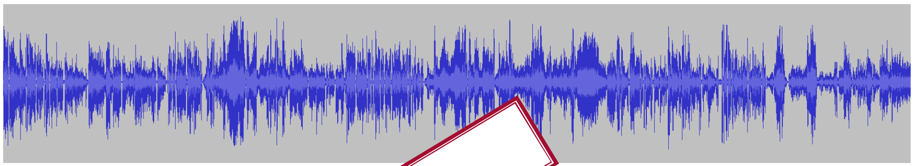
Royal Concertgebouw
Orchestra, Amsterdam

Mariss Jansons, conductor

* Arzt, A., Frostel, H., Gasser, M., Grachten, M. & Widmer, G. (2015). Artificial Intelligence in the Concertgebouw. In *Proceedings of the 24th International Joint Conference on Artificial Intelligence (IJCAI 2015)*, Buenos Aires, Argentina.

JKU

COMPUTATIONAL MUSIC PERCEPTION



'EASY'

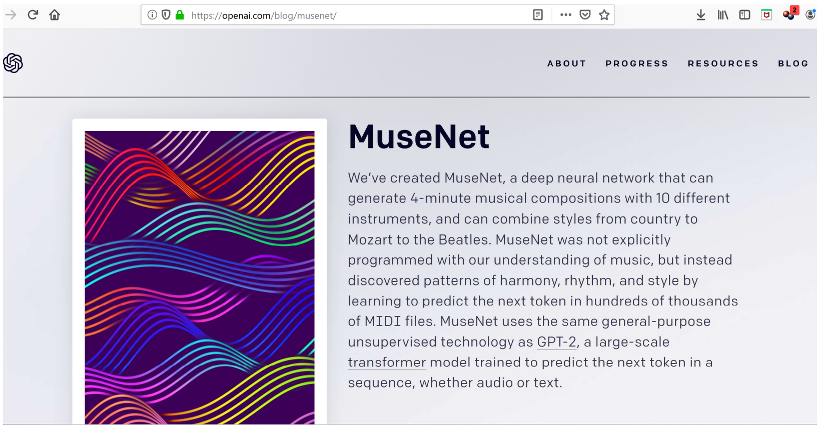
REAL-TIME IDENT... AND TRACKING
GENERAL MU... ILARITY (e.g., for recommendation)
GENRE, S... Z / Rock / Pop / Folk / HipHop / ...
SEGMENT S... UCTURE: Chorus, Verse, Repeats, ...
HARMONIC STRUCTURE: Key, Chords, Chord Sequences
RHYTHMIC STRUCTURE: Beat, Tempo, Measures/Bars, Rhythm
MUSIC DETECTION

JKU

THEN WHAT IS HARD?

JKU

COMPOSING NEW MUSIC



MuseNet / GPT-2 (2019)
https://openai.com/blog/musenet/

PLAYING MUSIC EXPRESSIVELY

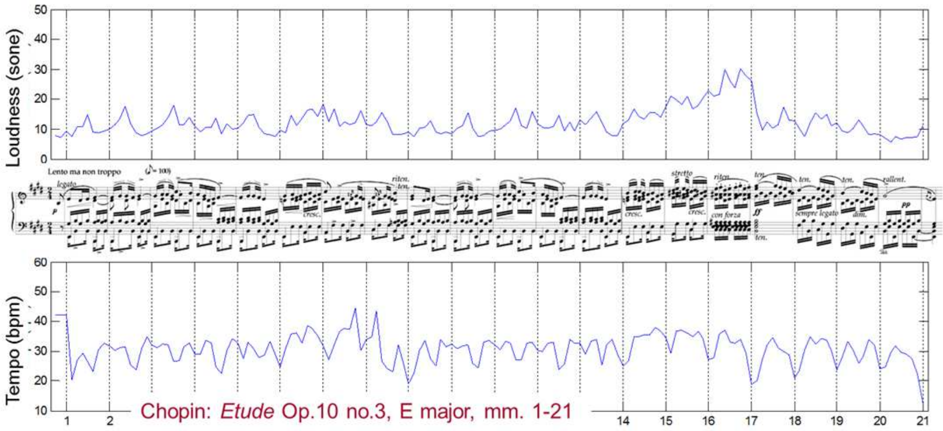
➔ **Play this piece!**

Arthur Rubinstein (1967):

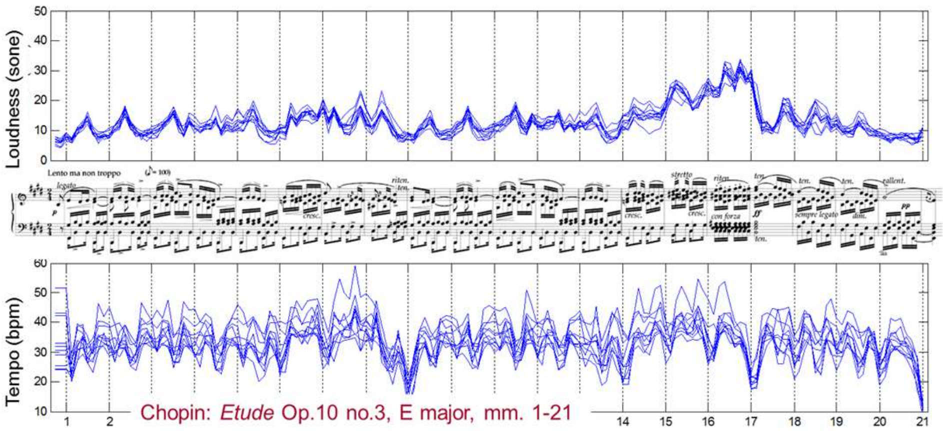
Nikita Magaloff (1989):

Frederic Chopin,
Nocturne Op. 9 No. 1
Bb minor

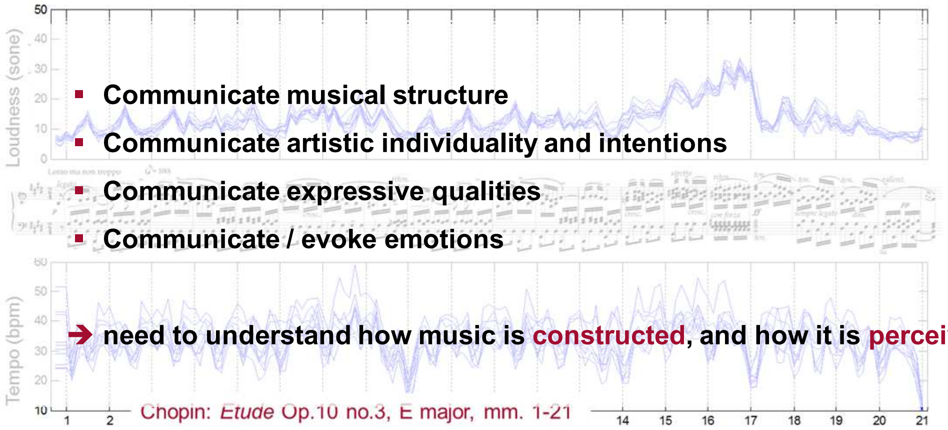
EXPRESSIVE MUSIC PERFORMANCE



EXPRESSIVE MUSIC PERFORMANCE



EXPRESSIVE MUSIC PERFORMANCE



- Communicate musical structure
- Communicate artistic individuality and intentions
- Communicate expressive qualities
- Communicate / evoke emotions

➔ need to understand how music is constructed, and how it is perceived

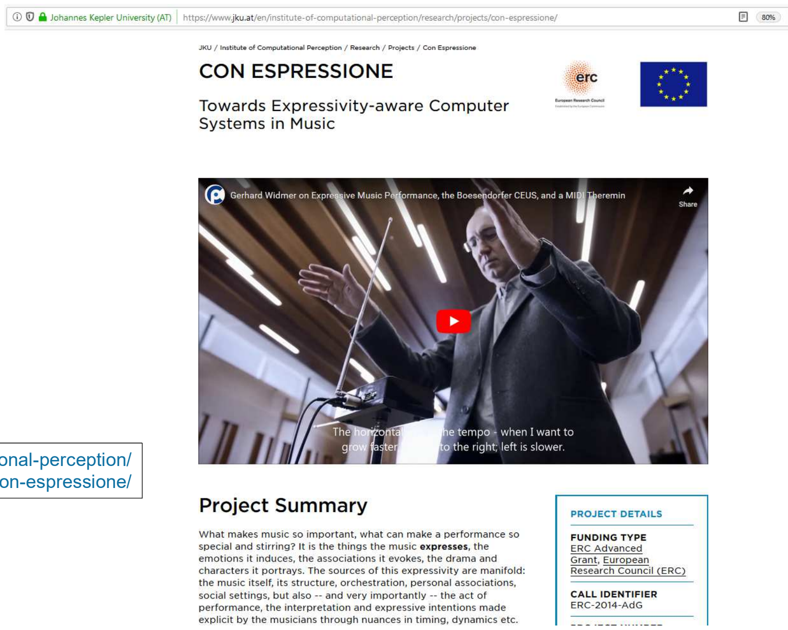
Chopin: *Etude* Op.10 no.3, E major, mm. 1-21

JKU

THE *CON ESPRESSIONE* PROJECT

European Research Council
Established by the European Commission

www.jku.at/en/institute-of-computational-perception/research/projects/con-espressione/



JKU

THE *CON ESPRESSIONE* PROJECT

European Research Council
Established by the European Commission

- **What kinds of expressive qualities do listeners perceive / differentiate?**
- **What is it in a performance that communicates an expressive quality?**
- **Can machines learn to *recognise* expressive qualities?**
- **Can machines learn to *play* music "expressively"?**
- **Can machines become truly 'musical' companions?**
- **… and what can we learn from all this?**

JKU

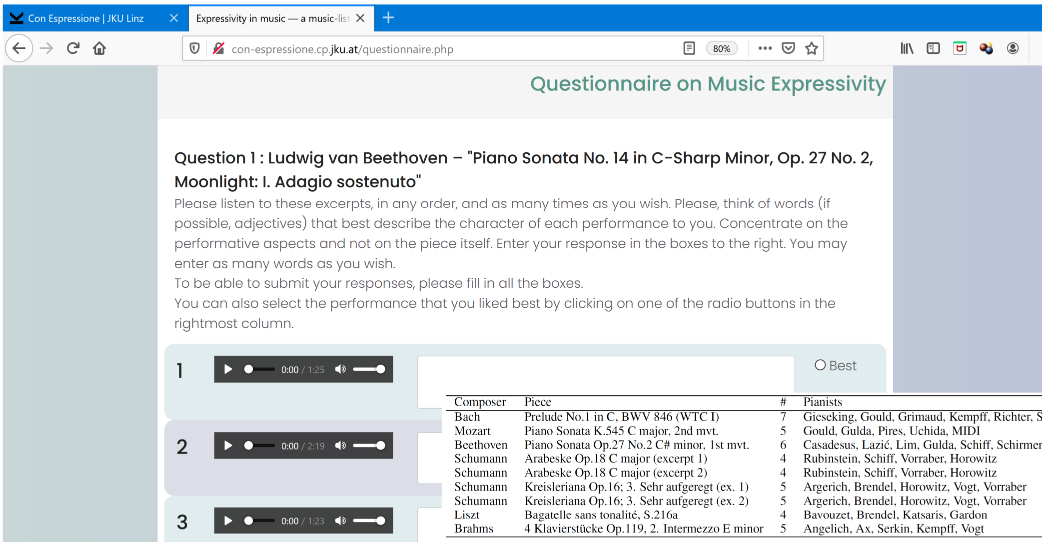
THE *CON ESPRESSIONE* PROJECT

European Research Council
Established by the European Commission

- **What kinds of expressive qualities do listeners perceive / differentiate?**
- What is it in a performance that communicates an expressive quality?
- Can machines learn to *recognise* expressive qualities?
- Can machines learn to *play* music "expressively"?
- Can machines become truly 'musical' companions?
- … and what can we learn from all this?

JKU

THE *CON ESPRESSIONE GAME*



Composer	Piece	#	Pianists
Bach	Prelude No.1 in C, BWV 846 (WTC I)	7	Gieseking, Gould, Grimaud, Kempff, Richter, Stadtfeld, MIDI
Mozart	Piano Sonata K.545 C major, 2nd mvt.	5	Gould, Gulda, Pires, Uchida, MIDI
Beethoven	Piano Sonata Op.27 No.2 C# minor, 1st mvt.	6	Casadesus, Lazić, Lim, Gulda, Schiff, Schirmer
Schumann	Arabeske Op.18 C major (excerpt 1)	4	Rubinstein, Schiff, Vorraber, Horowitz
Schumann	Arabeske Op.18 C major (excerpt 2)	4	Rubinstein, Schiff, Vorraber, Horowitz
Schumann	Kreisleriana Op.16; 3. Sehr aufgeregt (ex. 1)	5	Argerich, Brendel, Horowitz, Vogt, Vorraber
Schumann	Kreisleriana Op.16; 3. Sehr aufgeregt (ex. 2)	5	Argerich, Brendel, Horowitz, Vogt, Vorraber
Liszt	Bagatelle sans tonalité, S.216a	4	Bavouzet, Brendel, Katsaris, Gardon
Brahms	4 Klavierstücke Op.119, 2. Intermezzo E minor	5	Angelich, Ax, Serkin, Kempff, Vogt

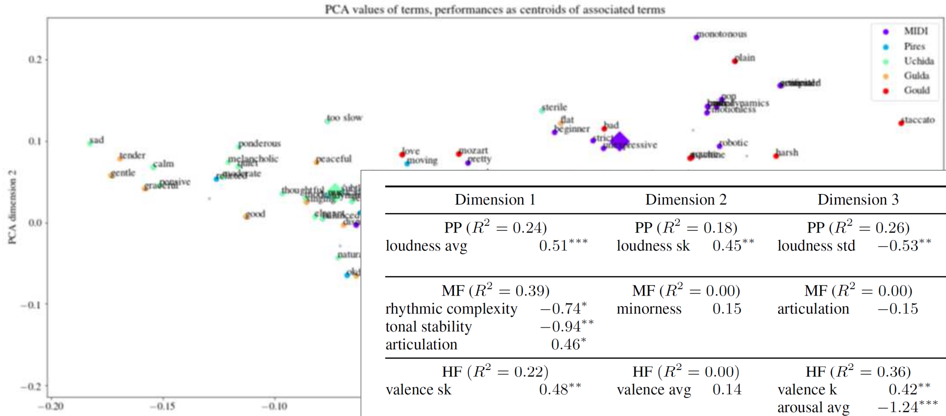
Table 1. Performances used in the *Con Espressione* Game.

THE *CON ESPRESSIONE GAME*

1.515 individual descriptions for a total of
3.166 terms, of which
1.415 are unique (approx. 45%)

* Cancino, C., Aljanaki, A., Chowdhury, S., Peter, S. and Widmer, G. (2020). On the Characterization of Expressive Performance in Classical Music: First Results of the Con Espressione Game. *Proc. ISMIR 2020*, Montreal, Canada.

THE *CON ESPRESSIONE* GAME



Dimension 1		Dimension 2		Dimension 3		Dimension 4	
PP ($R^2 = 0.24$)		PP ($R^2 = 0.18$)		PP ($R^2 = 0.26$)		PP ($R^2 = 0.24$)	
loudness avg	0.51^{***}	loudness sk	0.45^{**}	loudness std	-0.53^{**}	beat period k	-0.34^{*}
						loudness std	-0.44^{*}
MF ($R^2 = 0.39$)		MF ($R^2 = 0.00$)		MF ($R^2 = 0.00$)		MF ($R^2 = 0.29$)	
rhythmic complexity	-0.74^{*}	minorness	0.15	articulation	-0.15	rhythmic complexity	0.52^{*}
tonal stability	-0.94^{**}					tonal stability	0.84^{***}
articulation	0.46^{*}						
HF ($R^2 = 0.22$)		HF ($R^2 = 0.00$)		HF ($R^2 = 0.36$)		HF ($R^2 = 0.09$)	
valence sk	0.48^{**}	valence avg	0.14	valence k	0.42^{**}	valence k	-0.33^{*}
				arousal avg	-1.24^{***}		
				valence std	0.27^{*}		
				valence avg	-0.82^{*}		

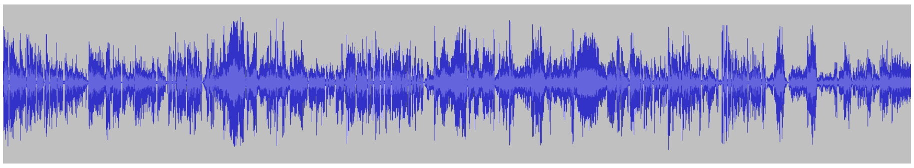
* Cancino, C., Aljanaki, A., Chowdhury, S., Peter, S. and Widmer, G. (2020). On the Characterization of Expressive Performance in Classical Music: First Results of the Con Espressione Game. *Proc. ISMIR 2020*, Montreal, Canada.

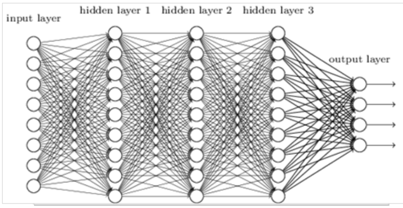
THE *CON ESPRESSIONE* PROJECT

European Research Council
Established by the European Commission

- What kinds of expressive qualities do listeners perceive / differentiate?
- **What is it in a performance that communicates an expressive quality?**
- **Can machines learn to *recognise* expressive qualities?**
- Can machines learn to *play* music "expressively"?
- Can machines become truly 'musical' companions?
- … and what can we learn from all this?

PREDICTING AND EXPLAINING EXPRESSIVE QUALITIES?

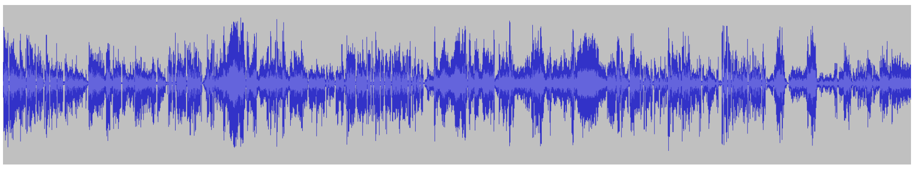


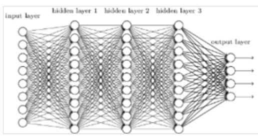


Perceived expressive qualities

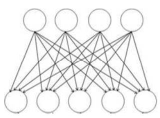
JKU

PREDICTING AND EXPLAINING EXPRESSIVE QUALITIES?





Intuitive musical percepts („mid-level feature



Perceived expressive qualities

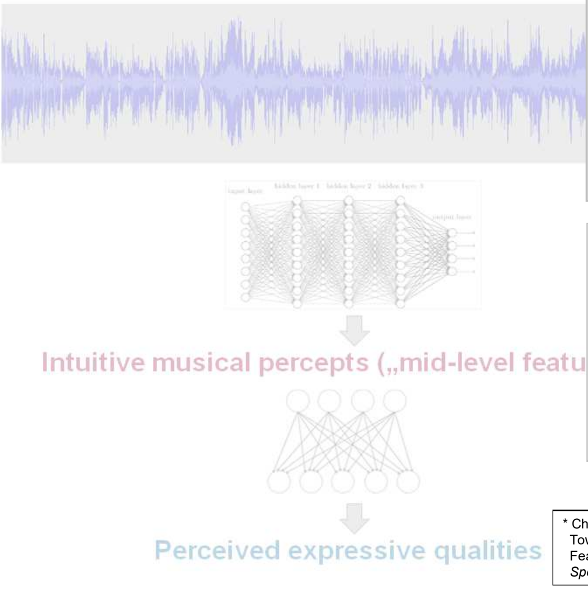
Perceptual Feature	Question asked to human raters
Melodiousness	To which excerpt do you feel like singing along?
Articulation	Which has more sounds with staccato articulation?
Rhythmic Stability	Imagine marching along with the music. Which is easier to march along with?
Rhythmic Complexity	Is it difficult to repeat by tapping? Is it difficult to find the meter? Does the rhythm have many layers?
Dissonance	Which excerpt has noisier timbre? Has more dissonant intervals (tritones, seconds, etc.)?
Tonal Stability	Where is it easier to determine the tonic and key? In which excerpt are there more modulations?
Modality ('Minorness')	Imagine accompanying this song with chords. Which song would have more minor chords?

[Aljanaki & Soleymani, ISMIR 2018]

* Haunschmid, V., Chowdhury, S. and Widmer, G. (2019). Two-level Explanations in Music Emotion Recognition. *ML4MD Workshop, 36th Intl. Conference on Machine Learning (ICML 2019)*, Long Beach, CA.

JKU

PREDICTING AND EXPLAINING EXPRESSIVE QUALITIES?



	Dim 1	Dim 2	Dim 3	Dim 4
VGG-ish	0.35	0.10	0.22	0.32
RF-ResNet	0.36	0.07	0.28	0.33
RF-ResNet DA	**0.40**	0.09	**0.29**	0.32
RF-ResNet DA+TS	0.35	**0.15**	**0.29**	**0.34**

Table 1. Coefficient of determination (R^2-score) of description embedding dimensions of the Con Espressione game using a linear regressor trained on predicted mid-level features.

RF-ResNet		RF-ResNet DA+TS	
Feature	r	Feature	r
articulation	0.47	melodiousness	− 0.39
rhythmic complexity	0.41	articulation	0.46
		rhythmic complexity	0.41
		dissonance	0.40

Table 2. Pearson's correlation (r) for mid-level features with the first description embedding dimension, with (right) and without (left) domain adaptation. Features with $p < 0.05$ and $|r| > 0.20$ are selected. This dimension has positive loadings for words like "hectic", "irregular", and negative loadings for words like "sad", "gentle", "tender".

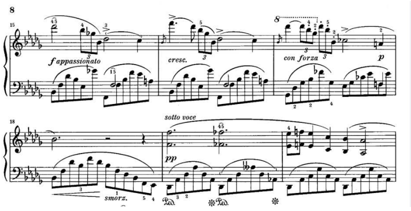
* Chowdhury, S. and Widmer, G. (2021).
Towards Explaining Expressive Qualities in Piano Recordings: Transfer of Explanatory Features via Acoustic Domain Adaptation. *Proc. 46th IEEE Int. Conf. on Acoustics, Speech & Signal Processing (ICASSP 2021), Toronto, Canada.*

THE *CON ESPRESSIONE* PROJECT

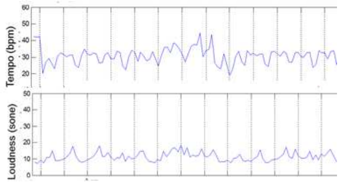
European Research Council
Established by the European Commission

- What kinds of expressive qualities do listeners perceive / differentiate?
- What is it in a performance that communicates an expressive quality?
- Can machines learn to *recognise* expressive qualities?
- **Can machines learn to *play* music "expressively"?**
- Can machines become truly 'musical' companions?
- … and what can we learn from all this?

COMPUTATIONAL MODELS OF EXPRESSIVE PERFORMANCE

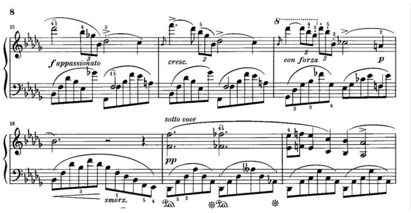


Predictive Model



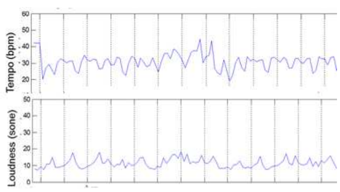
JKU

COMPUTATIONAL MODELS OF EXPRESSIVE PERFORMANCE



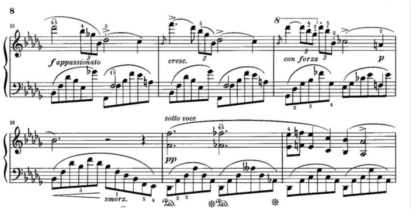
note	$\varphi_{stacc}(.)$	$\varphi_{grace}(.)$	$\varphi_{pitch}(.)$	$\varphi_{pitch^2}(.)$	$\varphi_{cresc}(.)$
x_1	1	0	$\frac{72}{127}$	$\left(\frac{72}{127}\right)^2$	0.00
x_2	0	0	$\frac{76}{127}$	$\left(\frac{76}{127}\right)^2$	0.25
x_3	0	0	$\frac{72}{127}$	$\left(\frac{72}{127}\right)^2$	0.25
x_4	0	0	$\frac{67}{127}$	$\left(\frac{67}{127}\right)^2$	0.25
x_5	0	1	$\frac{74}{127}$	$\left(\frac{74}{127}\right)^2$	0.75
x_6	0	0	$\frac{76}{127}$	$\left(\frac{76}{127}\right)^2$	0.75

Predictive Model



JKU

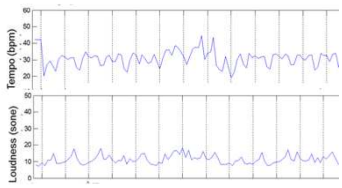
COMPUTATIONAL MODELS OF EXPRESSIVE PERFORMANCE



note	$\varphi_{stacc}(\cdot)$	$\varphi_{grace}(\cdot)$	$\varphi_{pitch}(\cdot)$	$\varphi_{pitch^2}(\cdot)$	$\varphi_{cresc}(\cdot)$
x_1	1	0	$\frac{72}{127}$	$\left(\frac{72}{127}\right)^2$	0.00
x_2	0	0	$\frac{76}{127}$	$\left(\frac{76}{127}\right)^2$	0.25
x_3	0	0	$\frac{72}{127}$	$\left(\frac{72}{127}\right)^2$	0.25
x_4	0	0	$\frac{67}{127}$	$\left(\frac{67}{127}\right)^2$	0.25
x_5	0	1	$\frac{74}{127}$	$\left(\frac{74}{127}\right)^2$	0.75
x_6	0	0	$\frac{76}{127}$	$\left(\frac{76}{127}\right)^2$	0.75

Linear mapping

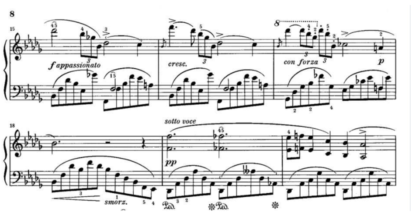
$$\hat{\boldsymbol{y}} = f(\boldsymbol{x}; \boldsymbol{w}) = \boldsymbol{w}^\top \boldsymbol{\varphi}(\boldsymbol{x})$$



JKU

* Cancino Chacón, C., Gadermaier, T., Widmer, G. and Grachten, M. (2017). An Evaluation of Linear and Non-linear Models of Expressive Dynamics in Classical Piano and Symphonic Music. *Machine Learning* 106(6), 887-909.

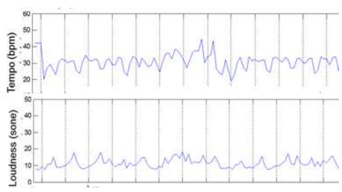
COMPUTATIONAL MODELS OF EXPRESSIVE PERFORMANCE



note	$\varphi_{stacc}(\cdot)$	$\varphi_{grace}(\cdot)$	$\varphi_{pitch}(\cdot)$	$\varphi_{pitch^2}(\cdot)$	$\varphi_{cresc}(\cdot)$
x_1	1	0	$\frac{72}{127}$	$\left(\frac{72}{127}\right)^2$	0.00
x_2	0	0	$\frac{76}{127}$	$\left(\frac{76}{127}\right)^2$	0.25
x_3	0	0	$\frac{72}{127}$	$\left(\frac{72}{127}\right)^2$	0.25
x_4	0	0	$\frac{67}{127}$	$\left(\frac{67}{127}\right)^2$	0.25
x_5	0	1	$\frac{74}{127}$	$\left(\frac{74}{127}\right)^2$	0.75
x_6	0	0	$\frac{76}{127}$	$\left(\frac{76}{127}\right)^2$	0.75

Non-linear mapping

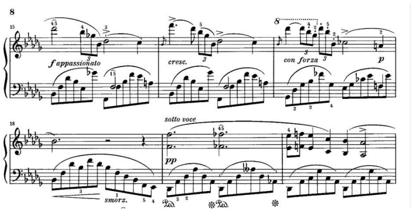
$$\begin{aligned} \hat{y}_i &= f^{(L)}\left(\boldsymbol{w}^{(L)\top}\boldsymbol{h}_i^{(L-1)} + w_0^{(L)}\right) \\ \boldsymbol{h}_i^{(l)} &= f^{(l)}\left(\boldsymbol{w}^{(l)\top}\boldsymbol{h}_i^{(l-1)} + \boldsymbol{w}_0^{(l)}\right) \\ \boldsymbol{h}_i^{(0)} &= \boldsymbol{\varphi}(x_i) \end{aligned}$$



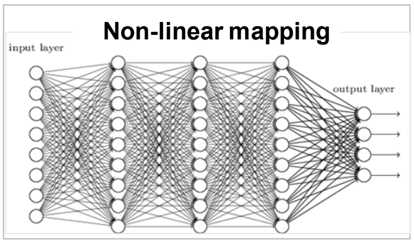
JKU

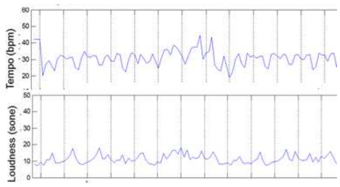
* Cancino Chacón, C., Gadermaier, T., Widmer, G. and Grachten, M. (2017). An Evaluation of Linear and Non-linear Models of Expressive Dynamics in Classical Piano and Symphonic Music. *Machine Learning* 106(6), 887-909.

COMPUTATIONAL MODELS OF EXPRESSIVE PERFORMANCE



note	$\varphi_{stacc}(\cdot)$	$\varphi_{grace}(\cdot)$	$\varphi_{pitch}(\cdot)$	$\varphi_{pitch^2}(\cdot)$	$\varphi_{cresc}(\cdot)$
x_1	1	0	$\frac{72}{127}$	$\left(\frac{72}{127}\right)^2$	0.00
x_2	0	0	$\frac{76}{127}$	$\left(\frac{76}{127}\right)^2$	0.25
x_3	0	0	$\frac{72}{127}$	$\left(\frac{72}{127}\right)^2$	0.25
x_4	0	0	$\frac{67}{127}$	$\left(\frac{67}{127}\right)^2$	0.25
x_5	0	1	$\frac{74}{127}$	$\left(\frac{74}{127}\right)^2$	0.75
x_6	0	0	$\frac{76}{127}$	$\left(\frac{76}{127}\right)^2$	0.75

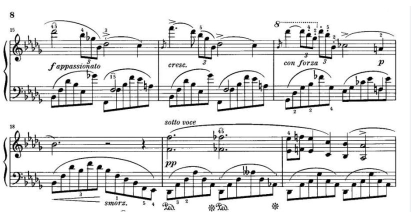




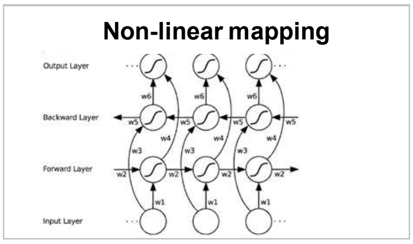
JKU

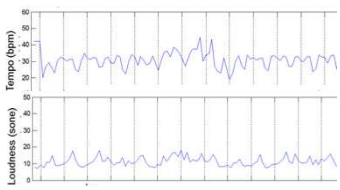
* Cancino Chacón, C., Gadermaier, T., Widmer, G. and Grachten, M. (2017). An Evaluation of Linear and Non-linear Models of Expressive Dynamics in Classical Piano and Symphonic Music. *Machine Learning* 106(6), 887-909.

COMPUTATIONAL MODELS OF EXPRESSIVE PERFORMANCE



note	$\varphi_{stacc}(\cdot)$	$\varphi_{grace}(\cdot)$	$\varphi_{pitch}(\cdot)$	$\varphi_{pitch^2}(\cdot)$	$\varphi_{cresc}(\cdot)$
x_1	1	0	$\frac{72}{127}$	$\left(\frac{72}{127}\right)^2$	0.00
x_2	0	0	$\frac{76}{127}$	$\left(\frac{76}{127}\right)^2$	0.25
x_3	0	0	$\frac{72}{127}$	$\left(\frac{72}{127}\right)^2$	0.25
x_4	0	0	$\frac{67}{127}$	$\left(\frac{67}{127}\right)^2$	0.25
x_5	0	1	$\frac{74}{127}$	$\left(\frac{74}{127}\right)^2$	0.75
x_6	0	0	$\frac{76}{127}$	$\left(\frac{76}{127}\right)^2$	0.75

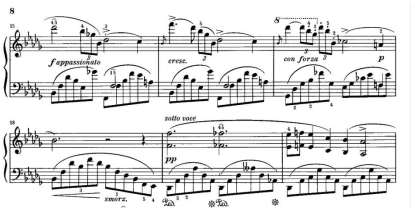




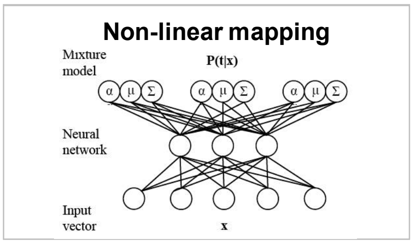
JKU

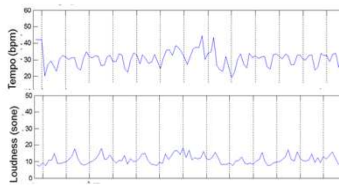
* Cancino Chacón, C., Gadermaier, T., Widmer, G. and Grachten, M. (2017). An Evaluation of Linear and Non-linear Models of Expressive Dynamics in Classical Piano and Symphonic Music. *Machine Learning* 106(6), 887-909.

COMPUTATIONAL MODELS OF EXPRESSIVE PERFORMANCE



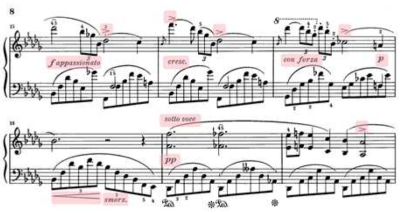
note	$\varphi_{stacc}(\cdot)$	$\varphi_{grace}(\cdot)$	$\varphi_{pitch}(\cdot)$	$\varphi_{pitch^2}(\cdot)$	$\varphi_{cresc}(\cdot)$
x_1	1	0	$\frac{72}{127}$	$\left(\frac{72}{127}\right)^2$	0.00
x_2	0	0	$\frac{76}{127}$	$\left(\frac{76}{127}\right)^2$	0.25
x_3	0	0	$\frac{72}{127}$	$\left(\frac{72}{127}\right)^2$	0.25
x_4	0	0	$\frac{67}{127}$	$\left(\frac{67}{127}\right)^2$	0.25
x_5	0	1	$\frac{74}{127}$	$\left(\frac{74}{127}\right)^2$	0.75
x_6	0	0	$\frac{76}{127}$	$\left(\frac{76}{127}\right)^2$	0.75

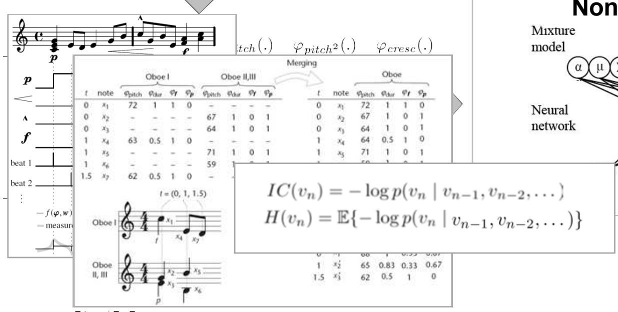


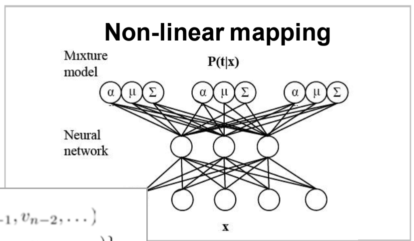


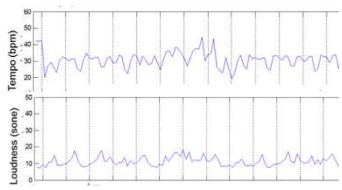
* Cancino Chacón, C., Gadermaier, T., Widmer, G. and Grachten, M. (2017). An Evaluation of Linear and Non-linear Models of Expressive Dynamics in Classical Piano and Symphonic Music. *Machine Learning* 106(6), 887-909.

COMPUTATIONAL MODELS OF EXPRESSIVE PERFORMANCE



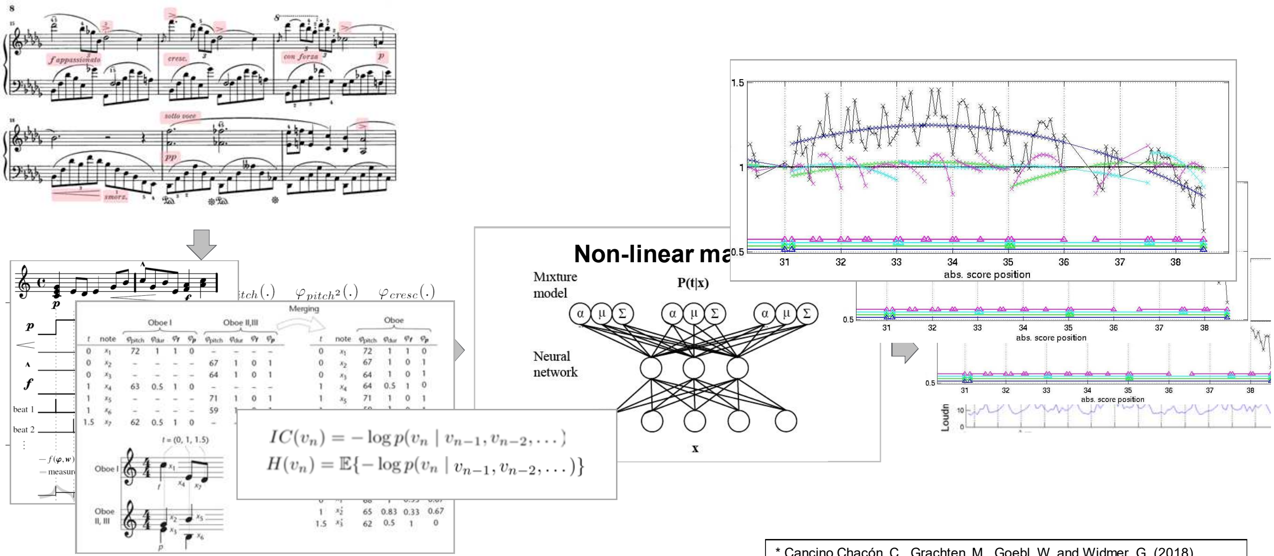






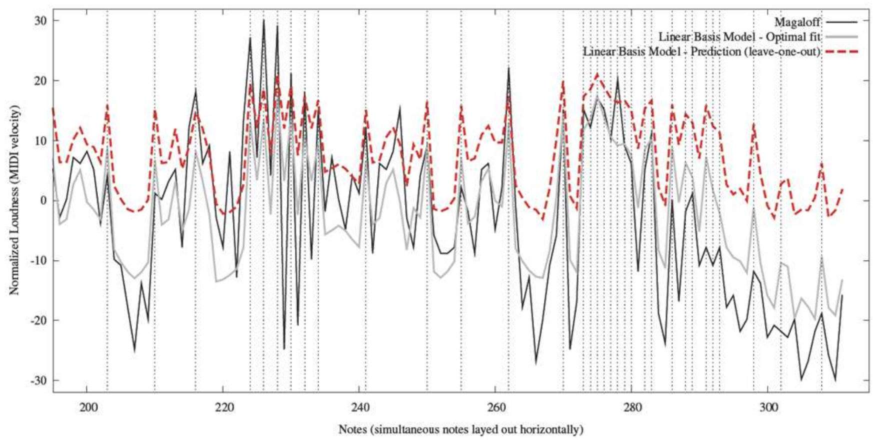
* Cancino Chacón, C., Grachten, M., Goebl, W. and Widmer, G. (2018). Computational Models of Expressive Music Performance: A Comprehensive and Critical Review. *Frontiers in Digital Humanities* 5:25 (2018).

COMPUTATIONAL MODELS OF EXPRESSIVE PERFORMANCE



* Cancino Chacón, C., Grachten, M., Goebl, W. and Widmer, G. (2018). Computational Models of Expressive Music Performance: A Comprehensive and Critical Review. *Frontiers in Digital Humanities* 5:25 (2018).

QUANTITATIVE EVALUATION



MUSICAL INSIGHTS

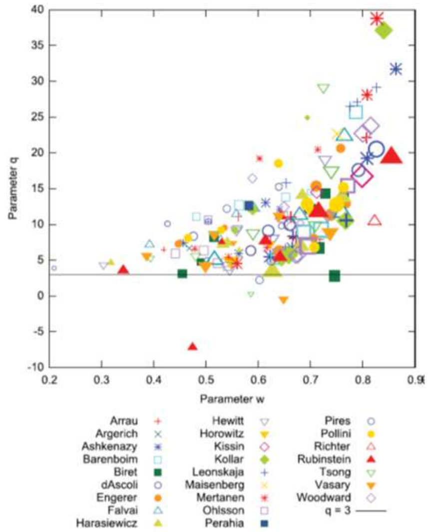


Figure 5. Distribution of rubato model parameter values over pianists; the size of the symbols is proportional to goodness of fit of the parameters to the data

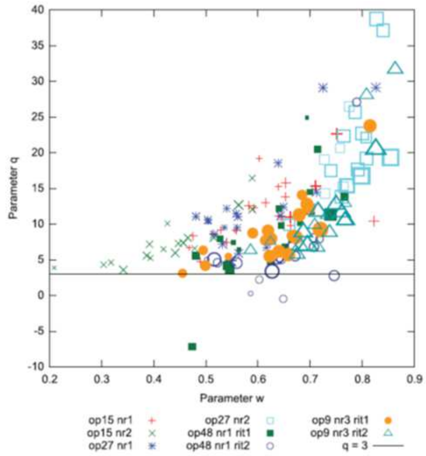
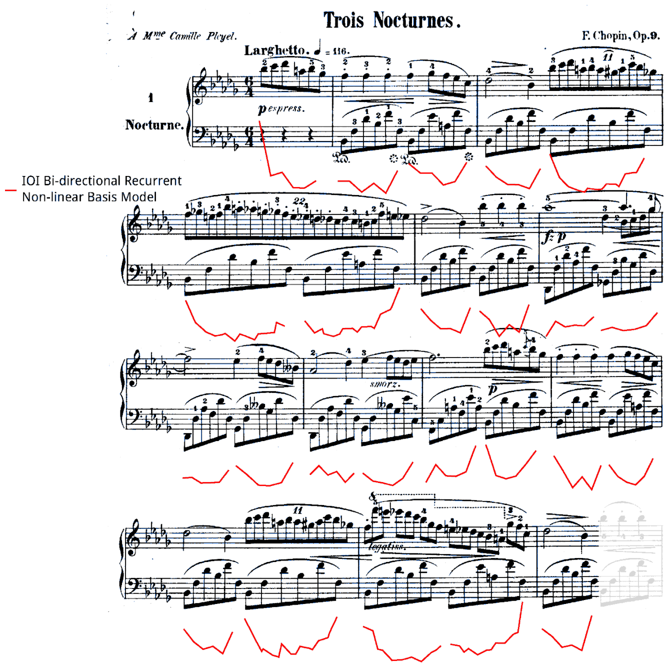


Figure 4. Distribution of rubato model parameter values over pieces; the size of the symbols is proportional to goodness of fit of the parameters to the data



Non-expressive literal performance:

The bi-directional recurrent non-linear model (2017):

A MUSICAL "TURING TEST" …

JKU

Journal of New Music Research, 2017
http://dx.doi.org/10.1080/09298215.2016.1264976

Routledge
Taylor & Francis Group

Algorithms can Mimic Human Piano Performance: The Deep Blues of Music

Emery Schubert[1], Sergio Canazza[2], Giovanni De Poli[2] and Antonio Rodà[2]

[1]UNSW Australia, Australia; [2]University of Padova, Italy

(*Received 18 July 2016; accepted 15 November 2016*)

Abstract

Can a computer play a music score, e.g. via a Disklavier, in a way that cannot be distinguished from a human performance of the same music? One hundred and seventy-two participants with a wide range of music playing backgrounds rated sound recordings of 7 performances of piano music by Kuhlau, one played by a human, and six generated by algorithms, including a 'mechanical' and an 'unmusical' rendering. Participants rated the extent to which each performance was by a human and explained their answers. The mechanical performance had the lowest mean rating, but the human performance was rated as statistically identical to the other stimuli. There were no differences between ratings made by classical piano experts and lay listeners, but despite this, the musicians were more confident with their ratings. Qualitative analysis revealed five broad themes that contribute to judging whether a piece appears to be human. The themes were labelled (in descending order of frequency) intuitive,

1. Introduction

Computer software has been able to send messages to acoustic pianos since 1980s allowing the development of automated performances of piano music on the piano (De Poli, 2004). With recent improvements in algorithmic generation of standard (mostly piano) classical/romantic repertoire, a new research question has been emerging: Will there ever be a time when a listener cannot distinguish between an algorithm performing a piece (for example, via a Yamaha Disklavier the Bösendorfer SE reproducing piano, Goebl & Bresin, 2003) versus a recording of an expert human performer (playing on the same device)? The ability of an algorithm or robot to be human-like has been a matter of fascinat[…] the possibilities of automation and robotics arose (i[…] of music, see Kapur, 2005). A famous example is a […] that could beat a world champion chess player. The […] 'Deep Blue' was able to achieve this milestone b[…] the world champion, Gary Kasparov, in an officia[…] ment in 1997, following several years of failure.

* Schubert, E., Canazza, S., De Poli, G. & Rodà , A. (2017). *Journal of New Music Research* 46(2), 175-186, 2017.

JK

2017: A MUSICAL "TURING TEST" …

The Piece:

- Friedrich Kuhlau, *Allegro Burlesco*, Op.88 No.30

The Contestants:

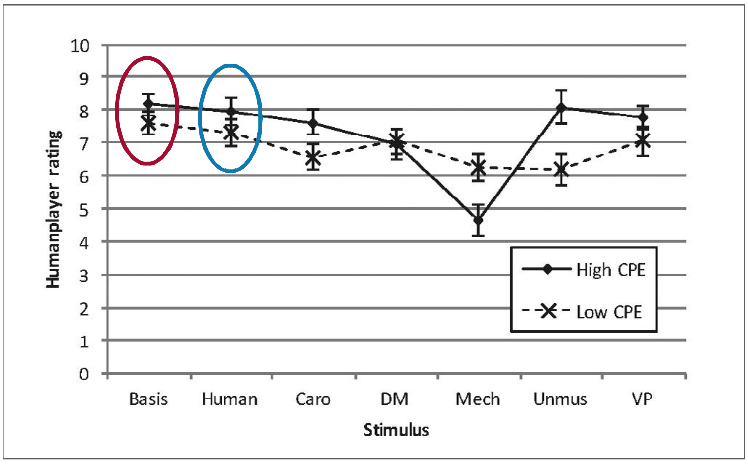
- 4 algorithms (CaRo, Director Musices, VirtualPhilharmony, Basis Function Model)
- 1 human "internationally renowned pianist" [Schubert *et al.,* 2017]
- 1 mechanical performance (deadpan)
- 1 "unmusical" performance (CaRo with inverted parameters)

The Evaluators:

- 172 listeners, different musical backgrounds, including pianists

JKU

2017: A MUSICAL "TURING TEST" …



"Humanplayer" rating by stimuli and expertise (mean and 1SE).
From (Schubert et al., *JNMR* 2017)

JKU

2017: A MUSICAL “TURING TEST” …

“This paper presents new evidence systematically demonstrating that algorithm-generated performanc[...] iano music can be indistinguishable from huma[...] es, suggesting some parallels with the 1990s vi[...] Deep Blue computer over the world ch[...] an) chess player.”

Really?

Schubert, E., Canazza, S., De Poli, G. and Rodà, A. (2017).
Algorithms Can Mimic Human Piano Performance: The Deep Blues of Music.
Journal of New Music Research 46(2), 175-186.

2017: A MUSICAL “TURING TEST” …

Friedrich Kuhlau, *Allegro Burlesco*, Op.88 No.30

(1)	(2)	(3)
Computer (mechanical)		

2017: A MUSICAL "TURING TEST" …

Friedrich Kuhlau, *Allegro Burlesco*, Op.88 No.30

(1)	(2)	(3)
Computer (mechanical)	**Computer (Basis Function Model)**	**Human Pianist**

JKU

THE *CON ESPRESSIONE* PROJECT

European Research Council
Established by the European Commission

- What kinds of expressive qualities do listeners perceive / differentiate?
- What is it in a performance that communicates an expressive quality?
- Can machines learn to *recognise* expressive qualities?
- Can machines learn to *play* music "expressively"?
- **Can machines become truly 'musical' companions?**
- … and what can we learn from all this?

JKU

EXPRESSIVE REACTIVE ACCOMPANIMENT: THE *ACCompanion*

Pianist:
Werner Goebl

(Sept. 2019)

JKU

THE *CON ESPRESSIONE* PROJECT

European Research Council
Established by the European Commission

- What kinds of expressive qualities do listeners perceive / differentiate?
- What is it in a performance that communicates an expressive quality?
- Can machines learn to *recognise* expressive qualities?
- Can machines learn to *play* music “expressively”?
- Can machines become truly ‘musical’ companions?
- **… and what can we learn from all this?**

JKU

THANK YOU!

This research receives funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme under grant agreement No 670035 (“Con Espressione”).

JOHANNES KEPLER UNIVERSITY LINZ