Classifying Leitmotifs in Recordings of Operas by Richard Wagner

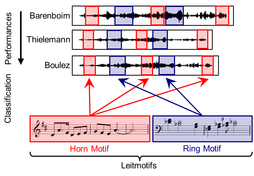
Michael Krause, Frank Zalkow, Julia Zalkow, Christof Weiß, Meinard Müller

1. Introduction

- **Leitmotifs**: short musical ideas associated with characters, places, items, or feelings
- Popular in 19th century **opera**, still used in contemporary **movie soundtracks**
- Major example: Richard Wagner's tetralogy *Der Ring des Nibelungen* (four operas, around 15 hours)
- This paper: Classifying leitmotif instances in music recordings

2. Leitmotif Classification

- Task: **Assigning a given audio excerpt to a class according to the leitmotif being played**



3. Experimental Dataset: Leitmotif Selection and Performances

- Our dataset: Based on the *Ring*, annotations publicly available
  - Selection of **ten central leitmotifs**
  - 16 recorded **performances**
  - 2403 motif **occurrences** in the score, annotated by a musicologist
  - Transfer of occurrences to performances (semi-automatically)

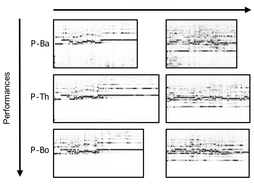
  ⟶ 16 × 2403 = 38448 **instances**!

Name	ID	Score	# Occurrences	Duration (Measures)
Nibelungen (Nibelungs)	L-Ni		536	0.96 ± 0.23
Ring (Ring)	L-Ri		286	1.49 ± 0.65
Mime (Mime)	L-Mi		242	0.83 ± 0.25
Nibelungenhass (Nibelungs' hate)	L-NH		237	0.96 ± 0.17
Ritt (Ride)	L-RT		228	0.66 ± 0.17
Waldweben (Forest murmurs)	L-Wa		223	1.10 ± 0.30
Waberlohe (Swirling blaze)	L-WL		190	1.21 ± 0.39
Horn (Horn)	L-Ho		172	1.38 ± 1.05
Geschwisterliebe (Siblings' love)	L-Ge		155	1.31 ± 0.83
Schwert (Sword)	L-Sc		134	1.89 ± 0.55

ID	Conductor	Year	hh:mm:ss
P-Ba	Barenboim	1991–92	14:54:55
P-Ha	Haitink	1988–91	14:27:10
P-Ka	Karajan	1967–70	14:58:08
P-Sa	Sawallisch	1989	14:06:50
P-So	Solti	1958–65	14:36:58
P-We	Weigle	2010–12	14:48:46
P-Bo	Boulez	1980–81	13:44:38
P-Böh	Böhm	1967–71	13:39:28
P-Fu	Furtwängler	1953	15:04:22
P-Ja	Janowski	1980–83	14:08:34
P-Ke	Keilberth/Furtwängler	1952–54	14:19:56
P-Kr	Krauss	1953	14:12:27
P-Le	Levine	1987–89	15:21:52
P-Ne	Neuhold	1993–95	14:04:35
P-Sw	Swarowsky	1968	14:56:34
P-Th	Thielemann	2011	14:31:13

4. Data Splits

- Two data splits:
  - *Performance split*: Requires generalizing to new accoustic and interpretation conditions
  - *Occurrence split*: Requires generalizing to new motif variations



5. LSTM-based Approach

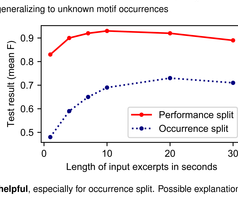
Layer	Output Shape	Parameters
Input	(V, 84)	
LSTM	(V, 84)	109056
LSTM	(V, 128)	131584
LSTM	(V, 128)	131584
Take last	(128)	
Batch normalization	(128)	512
Dense	(10)	1290
Output: Softmax	(10)	

- Input: **CQT** representations of audio excerpts
- Output: **Probabilities** per motif

6. Results

Split	Performance			Occurrence		
	P	R	F	P	R	F
L-Ni	0.94	0.95	0.94	0.67	0.80	0.73
L-Ri	0.93	0.92	0.93	0.36	0.41	0.38
L-Mi	0.96	0.95	0.96	0.79	0.87	0.83
L-NH	0.94	0.92	0.93	0.72	0.20	0.31
L-RT	0.95	0.94	0.95	0.57	0.65	0.61
L-Wa	0.94	0.98	0.96	0.87	0.80	0.84
L-WL	0.98	0.93	0.96	0.25	0.21	0.23
L-Ho	0.90	0.89	0.89	0.46	0.57	0.51
L-Ge	0.94	0.94	0.94	0.28	0.30	0.29
L-Sc	0.91	0.96	0.93	0.52	0.50	0.51
Mean	0.94	0.94	0.94	0.55	0.53	0.52

- Strong results when generalizing to unseen performances of the *Ring*
- Weaker results for generalizing to unknown motif occurrences



- **Temporal context helpful**, especially for occurrence split. Possible explanations:
  - Context encourages learning of *relevant musical characteristics* of motifs
  - Context increases possibility for *overfitting*
- Further evidence for overfitting: Classifier is able to **memorize** random labeling of occurrences

7. Future Work

- Next step: **Detecting** leitmotifs in entire performances (no presegmented inputs)
- Check out the **paper** and the accompanying **website**!
  https://www.audiolabs-erlangen.de/resources/MIR/2020-ISMIR-LeitmotifClassification

Acknowledgments

We thank Vlora Arifi-Müller for her assistance in preparing the data. This work was supported by the German Research Foundation (DFG MU 2686/7-2, MU 2686/11-1). The International Audio Laboratories Erlangen are a joint institution of the Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU) and Fraunhofer Institut für Integrierte Schaltungen IIS.

References

[1] Frank Zalkow, Christof Weiß, and Meinard Müller, "Exploring tonal-dramatic relationships in Richard Wagner's Ring cycle," in *Proceedings of the International Conference on Music Information Retrieval (ISMIR)*, Suzhou, China, 2017, pp. 642–648.

[2] Matthew Brabtzer-Stull, *Understanding the Leitmotif*, Cambridge University Press, 2015.

[3] Frank Zalkow, Christof Weiß, Thomas Prätzlich, Vlora Arifi-Müller, and Meinard Müller, "A multi-version approach for transferring measure annotations between music recordings," in *Proceedings of the AES International Conference on Semantic Audio*, Erlangen, Germany, 2017, pp. 148–155.

[4] Chiyuan Zhang, Samy Bengio, Moritz Hardt, Benjamin Recht, and Oriol Vinyals, "Understanding deep learning requires rethinking generalization," in *Proceedings of the International Conference on Learning Representations (ICLR)*, Toulon, France, 2017.