AUTOMATIC RECOGNITION OF INSTRUMENT FAMILIES IN POLYPHONIC RECORDINGS OF CLASSICAL MUSIC

Jakob Abeßer[1] **Christof Weiß**[1,2]

[1] Fraunhofer IDMT, Ilmenau, Germany

[2] International Audio Laboratories, Erlangen, Germany

jakob.abesser@idmt.fraunhofer.de

ABSTRACT

For analyzing Western classical music, instrumentation constitutes an important parameter. In a typical orchestral setting, we find several families comprising instruments with similar timbral characteristics. We therefore redefine the task of instrument recognition to the identification of the instrument families woodwinds, brass, piano, strings, and vocal. Our system relies on note events and is trained on monotimbral excerpts. In this paper, we visualize recognized note events together with score-based annotations for multitimbral examples from Western classical music.

1. INTRODUCTION

For classical music listeners, instrumental categories are of major importance. Some favor choir music, others prefer strings or piano. Regarding orchestral works, the instrumentation is often connected with the musical form, complexity, and expressive quality. Especially for long pieces such as symphonies, the variety of instrumental sounds provides opportunities for formal design and diverse affects. In operas, the choice of particular instruments and instrument groups may serve to emphasize the storyline or the emotional states of the characters. To study such kind of instrument usage, automatic analysis of instrumental activity throughout a piece can be helpful.

Publications on automatic instrument recognition (AIR) mainly focussed on classifying musical instruments based on *isolated note recordings* or *monophonic melodies*. For this tasks, the best systems were presented by Grasis et al. [5] and Tjoa and Liu [7] which achieved a mean class accuracy of 91.0 % for 10 instruments and 92.3 % for 24 instruments, respectively.

The analysis of *polyphonic music recordings* with one or multiple instruments playing is more relevant for practical MIR applications such as music recommendation and classification. In polyphonic music recordings, instrument sounds overlap and often share partial frequencies due to consonant interval relationships. To cope with this challenge, Barbedo and Tzanetakis [1] propose to *classify individual partials* instead of note events. Since the different partials often show very similar temporal envelopes, potentially overlapped partials can be excluded for the classification without losing information. Similarly, other authors propose to exclude spectral regions with strong overlap in the classification process [2, 4].

A different approach is to apply methods from research fields such as *source separation* and *automatic music transcription*. Heittola et al. [6] use the results of a multipitch estimation algorithm to perform source separation before classifying instruments. Similarly, Giannoulis et al. [4] combine multipitch estimation with AIR. In [3], Essid et al. focus on the likeliest instrument combination within a particular music genre. Based on extracted features, they group similar sounding instrumentations to an *automatically learned taxonomy* using hierarchical clustering.

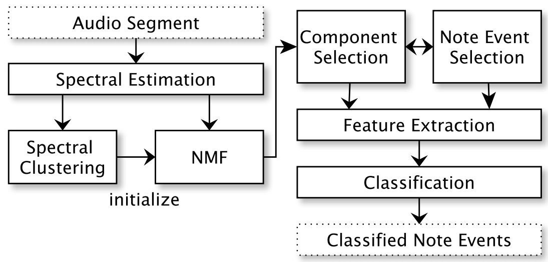

Figure 1: Processing flowchart of the proposed algorithm.

2. CLASSIFICATION OF MONOTIMBRAL MUSIC

This paper presents a new method for automatic recognition of instrument families in polyphonic recordings of classical music. Figure 1 shows the processing flow. For training, we use a set of monophonic examples for each class. We decompose the audio signals using Non-Negative Matrix Factorization (NMF) initialized by spectral clustering. From this, we retrieve note events by segmenting the temporal activations of the NMF. Therefore, our method does not require any prior transcription of the audio recording. Next, we select those NMF components that likely resemble monophonic note events. With each note event being represented by its spectral template and temporal envelope (including its derivative), we extract

© Jakob Abeßer, Christof Weiß. Licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0).

Attribution: Jakob Abeßer, Christof Weiß. "Automatic Recognition of Instrument Families in Polyphonic Recordings of Classical Music", Extended abstracts for the Late-Breaking Demo Session of the 16th International Society for Music Information Retrieval Conference, 2015.

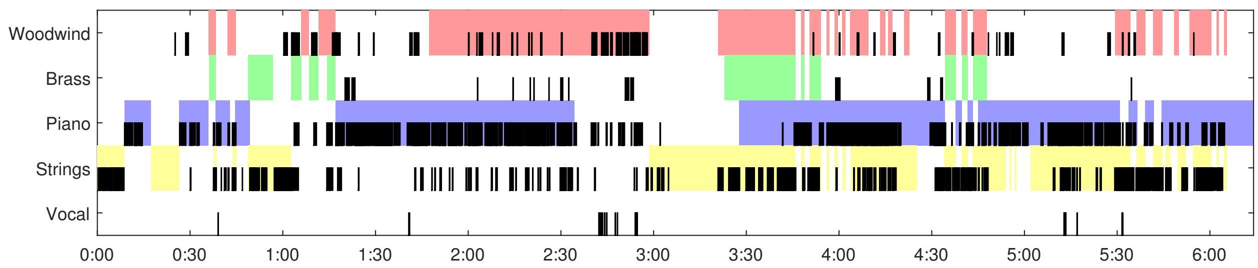

Figure 2: Classification results of a multiclass classifier for Beethoven's "Fantasia in C Minor" op. 80 (Choral Fantasy), first part of the Finale. The x-axis shows time in minutes. For each instrument family class, the reference instrument activity segments are given as colored rectangles. Note events classified towards each class are given as black lines.

	wood	*brass*	*piano*	*strings*	*vocal*
wood	**93.26**	3.37	2.25	1.12	0.00
brass	26.39	**36.11**	21.30	14.35	1.85
piano	0.45	0.22	**96.66**	1.56	1.11
strings	1.12	0.00	2.97	**92.94**	2.97
vocal	5.26	0.00	0.58	11.11	**83.04**

Table 1: Confusion matrix for classifying monotimbral audio excerpts via majority voting. The weighted mean F-measure is $\overline{F} = 83.3$. All values are given in percent.

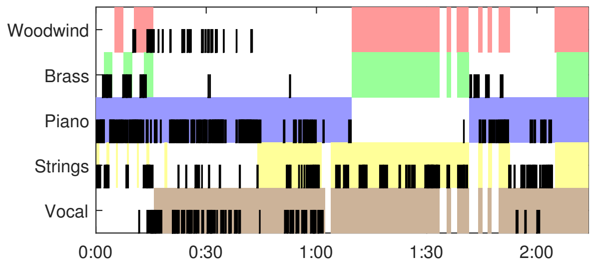

Figure 3: Multiclass classification results for the *Allegretto* part (Finale) from Beethoven's Choral Fantasy.

timbre-related features on both representations. Finally, we use the ExtraTreesClassifier to distinguish between instrument families. Table 1 shows the confusion matrix of a 10-fold cross validation over our dataset of monotimbral recordings. We perform a majority voting over 10 s segments. The results are promising. However, major problems occur for the brass family. One reason for this may be the french horn which is part of woodwind ensembles.

3. APPLICATION TO MULTITIMBRAL MUSIC

With the classifier trained on the five instrument families, we analyze the activity of the classes on a multi-instrumental polyphonic music recording. We choose the second part (Finale) of Beethoven's "Choral Fantasy" op. 80 for choir, soli, piano, and orchestra. Figures 2 and 3 visualize the results. The annotated instrument family activities are indicated by colored rectangles. The black lines show the note events classified to each instrument family. Comparing the detected events with the ground truth annotations, we generally find promising results especially for sections with only one or two active instrument groups. For more complex mixtures, training data with mixed groups may be useful. Just as for the classification of monotimbral segments, major problems occur for the classes woodwind and brass. This may arise due to the french horn or from an insufficient modeling of instrument characteristics.

4. CONCLUSION

We have presented a novel approach for automatic recognition of instrument families in polyphonic recordings of classical music. The algorithm does not require an initial music transcription step but relies on estimated note event candidates based on NMF decomposition results. Apart from the brass family, the system achieves high recognition rates for classifying instrument families. Finally, we have tested the system for multitimbral music. The algorithm performs well for small numbers of simultaneously active instrument families. Future work must incorporate strategies to cope with higher degrees of polyphony.

5. REFERENCES

[1] Jayme G. A. Barbedo and George Tzanetakis. Musical Instrument Classification Using Individual Partials. *IEEE Transactions on Audio, Speech, and Language Processing*, 19(1):111–122, 2011.

[2] Jana Eggink and Guy J. Brown. A Missing Feature Approach to Instrument Identification in Polyphonic Music. In *Proceedings of the IEEE Int. Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pages 553–556, Hong Kong and China, 2003.

[3] Slim Essid, Gael Richard, and Bertrand David. Instrument Recognition in Polyphonic Music Based on Automatic Taxonomies. *IEEE Transactions on Audio, Speech, and Language Processing*, 14(1):68–80, 2006.

[4] D. Giannoulis, E. Benetos, A. Klapuri, and M. D. Plumbley. Improving instrument recognition in polyphonic music through system integration. In *IEEE Int. Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 5222–5226, 2014.

[5] Mikus Grasis, Jakob Abeßer, Christian Dittmar, and Hanna Lukashevich. A Multiple-Expert Framework for Instrument Recognition. In *Proceedings of the 10th Int. Symposium on Computer Music Modelling and Retrieval (CMMR)*, 2013.

[6] Toni Heittola, Anssi Klapuri, and Tuomas Virtanen. Musical Instrument Recognition in Polyphonic Audio using Source-Filter Model for Sound Separation. In *Proceedings of the 10th Int. Society for Music Information Retrieval Conference (ISMIR)*, Kobe, Japan, 2009.

[7] Steven K. Tjao and K. J. Ray Liu. Musical Instrument Recognition using Biologically Inspired Filtering of Temporal Dictionary Atoms. In *Proceedings of the 11th Int. Society for Music Information Retrieval Conference (ISMIR)*, pages 435–440, 2010.