Introducing the Jazzomat project - Jazz solo analysis using Music Information Retrieval methods

Jakob Abeßer[1,2], Klaus Frieler[1], Martin Pfleiderer[1], and Wolf-Georg Zaddach[1]

[1]The Liszt School of Music, Weimar, Germany
[2]Semantic Music Technologies Group, Fraunhofer IDMT, Ilmenau, Germany

Abstract. The JAZZOMAT project investigates the creative processes underlying jazz solo improvisations with the help of statistical and computational methods. To this end, a comprehensive and representative database of jazz solo transcription is being built up, and an open-source Python Library is developed for analysis purposes. Besides the general outline of the project and a description of our core feature module, we present three typical analysis tasks along with some preliminary and exemplary results.

Keywords: Jazz, improvisation, computational musicology, pattern mining

1 Introduction

The main focus of the JAZZOMAT project lies on the statistical analysis of monophonic jazz solos with the overarching goal to explore the cognitive and cultural foundations of jazz solo improvisation. The project has a strong interdisciplinary approach combining jazz research, cognitive music psychology, computational (ethno-)musicology, as well as music information retrieval (MIR) [1].

Computational and statistical analysis of (folk) songs have a long-standing tradition in ethnomusicology (see [2] for a historical overview), but during the last years it also gained momentum in music psychology (e. g. [3,4,5]), because statistical regularities are thought to play a important part in music cognition and might also provide insights into the underlying creative processes. For instance, one theory of jazz improvisation postulates the existence of a set of personal licks and formulas for improvisers. Testing this theory is a pattern mining task, a standard problem in MIR. However, the definition of cognitively adequate or instrument-specific patterns is still an open question. Previous appoaches (e.g., [6,7,8]) use a variety of definitions, but no general consensus has been achieved.

On the other hand, several more theories for jazz improvisation exist, but evidence is mostly based on single case studies (e. g., [9]), and they remain largely unverified on a representative data corpus (cf. [10]). One notable exception is the work of Thomas Owens [11], who was able to catalog 193 recurring melodic patterns (formulas) in a comprehensive set of 250 Charlie Parker solo transcriptions. As a follow-up, Pierre-Yves Rolland [12] devised a pattern mining algorithm which allegedly reproduced most of Owens's formulas but also found previously undiscovered ones, although no detailed results were published.

To address these and other questions, we see the need for a comprehensive (digital) database of jazz solo transcriptions as a solid foundation for theory testing, as well as for flexible and powerful tools to carry out the necessary analyses and computations.

2 Project Components

2.1 Database of Jazz Solo Transcriptions

A central goal of the JAZZOMAT project is the construction of a novel database of high-quality jazz solo transcriptions, which will be made public in the near future. State-of-the-art MIR tools are used such as *Songs2See*[1], which is used to automatically extract and transcribe solos directly from recordings, and *SmartScore*[2], which is used to scan and digitalize existing sheet music taken from the literature or other public sources. Both the automatic and the existing transcriptions are often erroneous or incomplete. Therefore, they are cross-checked and corrected by jazz and musicology students at the Liszt School of Music with the help of *SonicVisualiser*[3]. Additionally, various metadata of a given solo part are annotated:

– phrases (grouping of melody notes),
– beats (manually tapped),
– chords (taken from the sheet music),
– musical form (e.g. AABA, 12-bar blues),
– articulation.

The focus of the transcription lies on the syntactic and not on the expressive and semantic level. Improvisations are encoded as metrical and harmonically annotated note sequences with the parameters pitch, onset, and duration. Loudness and timbre (e.g. playing techniques) information are currently not considered but might be included in the future.

2.2 MeloSpy - a Modular Analysis Framework in Python

Within the JAZZOMAT project, a novel Python-based analysis framework called *MeloSpy* is currently under development. This modular framework allows to analyze arbitrary monophonic symbolic music using different features. These features can be defined by the user by means of the `melfeature` component (aka the "feature machine").

`melfeature`. In the field of MIR, transcription-based (high-level) features often consist of very basic statistical measures of note-event parameters such as pitch, onset, inter-onset-interval, or duration (cf., [13]). Following and extending this concept, `melfeature` was designed to combine multiple analysis modules (such as histogram computation or statistical measures like mean or variance) to module chains. Modules

[1] `http://www.songs2see.com/`
[2] `http://www.klemm-music.de/musitek/smartscore/`
[3] `http://www.sonicvisualiser.org`

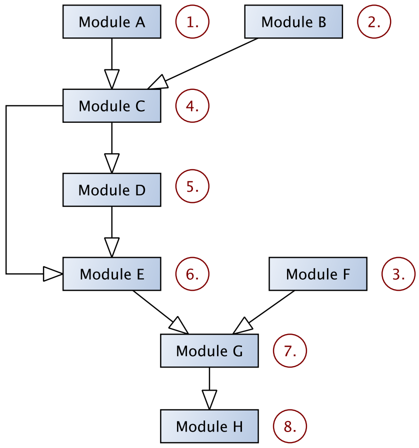

Fig. 1. Example feature configuration with 8 modules and 8 module connectors. The processing order is shown in red font.

can be grouped arbitrarily in a serial or parallel fashion as shown in Figure 1. The processing order of all modules is determined automatically by resolving the dependencies between the modules. A processing chain of modules will be hereafter called *feature*. A feature's output is not restricted to single numbers, but can be also vectors of complex data structures.

Feature definition with configuration files. Features are defined using configuration files in the `YAML` language, which are easy readable and writable. Example 1 illustrates a simple feature—the pitch range (ambitus) in semi-tones of a given note sequence.

Example 1

```
label: pitch_range
description: Pitch range in semitones
feature:
  source:
    PITCH:
      param: pitch
  process:
    STAT:
      type: stat
      measure: range
      inputVec: PITCH.outputVec
  sink:
    PITCH_RANGE_FEATURE:
      input: STAT.outputVec
      label: pitch_range
```

Within the configuration files, a number of `source` modules allow to access different note event parameters such as pitch, interval, fuzzy interval[4], duration in beats, inter-onset interval, or Parson code[5] from a given melody.

In the `process` section, the feature's core functionality is implemented by grouping different modules using input-output-connectors. `melfeature` provides the user with a large number of basic module types such as modules for descriptive statistics and histograms, logical and arithmetic operations, selectors, or n-grams of arbitrary vectors. Each module has a specific set of mandatory and optional input and output parameters.

Finally, in the `sink` section, different process chain sinks are defined that receive the output of a processing module and save it as a feature for further analysis. This way, an arbitrary amount of features can be computed for a list of files.

Amongst others, the current list of features includes histograms, interval distributions, n-gram computation, event density, tempo, intra-phrase similarity, chromaticism, and a contour segmentation.

Temporal aggregation within features. `melfeature` allows to compute feature values over individual structural elements of a melody such as phrases, form parts, or choruses. These features values can then be aggregated using different statistical measures to obtain a final feature value.

Batch processing. `melfeature` will include a large set of pre-defined features tailored towards the note event representations used in the JAZZOMAT project. Batch processing of multiple SonicVisualizer files using multiple feature configuration files is supported. Additionally, batch processing can be performed separately on structural elements of a melody such as phrases, form parts, or choruses for each solo (without aggregation).

3 Analysis Tasks & Preliminary Results

In the following section, three different analysis tasks will be briefly discussed and some preliminary results presented to illustrate how MIR methods can support musicological research in our project.

3.1 Retrieval of repeating patterns

Automated pattern retrieval in the symbolic domain is a popular research field in MIR (e.g, [14] for an overview). Most methods aim to detect exact repetitions (possibly with respect to an abstract and/or reduced representation) in monophonic or polyphonic music. However, in jazz solos, patterns are often repeated with slight variations [12], which adds extra challenges to the already highly combinatorial problem. This problem will be addressed in the future, at the current stage of the project, we retrieve only exact

[4] The fuzzy interval representation groups related intervals such as minor and major thirds in the same class.

[5] The Parson code (or contour) encodes the interval direction.

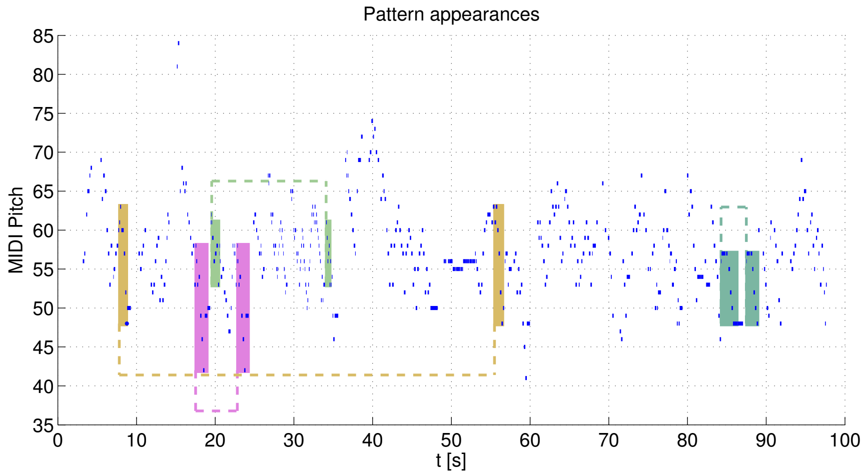

Fig. 2. Pairs of exact pitch patterns in Clifford Brown's solo on "Joy Spring". Notes are indicated as piano roll (blue rectangles), pattern appearances as colored rectangles connected by dotted lines.

pattern repetitions on note dimensions pitch and interval from diagonal paths in the self-similarity matrix. Figure 2 illustrates occurrences of exact pitch patterns (with a minimum pattern length of 6) in Clifford Brown's solo on "Joy Spring"[6] . The distance between pattern occurrences indicates if a pattern is part of a motivic sequence (e.g., the olive pattern at about 86–88s in the solo) or a more generic patttern from the formula repertoire of the soloist (e.g. the brown pattern at 8s and 56s).

As a next step, we aim to group similar patterns using clustering methods and to determine pattern prototypes for clusters in order to reconstruct underlying cognitive templates. Furthermore, we will search for patterns that occur in different solos of the same artist to get a deeper insight into the artist's personal repertoire, and for patterns that occur across players and thus might be transmitted by imitation.

3.2 Phrase similarity

Phrases are the basic building blocks of melodies. They are groupings of notes which are perceived (and produced) as a single musical unit. It is therefore crucial for our analysis purposes that our solo transcriptions contain phrase annotations.

A fairly general construction principle in melodic invention is the concept of varied repetition [15], i.e., a sequence of repeated but instantly varied melodic units. In order to automatically detect phrase groups that follow this construction principle, we compute the similarity matrix over all phrases using similarity measure that is based on the edit distance of pitch. The similarity matrix over all phrases of John Coltrane's solo on "So What"[7] is shown in Figure 3. The upper-side diagonal can be analyzed for high

[6] Rec. 1945, Nippon Phonogram Co., Ldt0 838 308-2
[7] Rec. 1959, CBS 460603 2

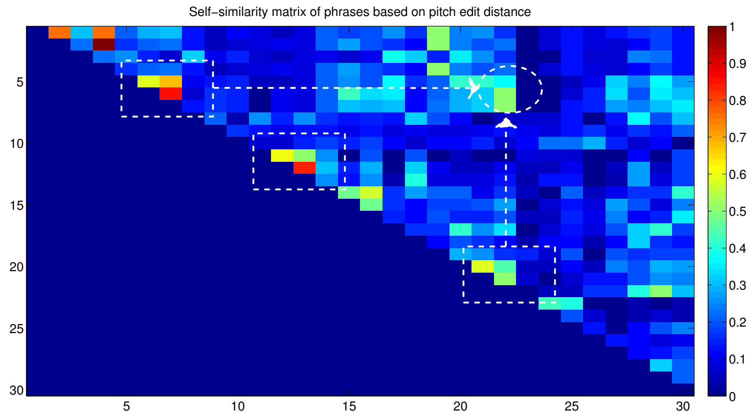

Fig. 3. Edit distance based similarity between different phrases of John Coltrane's solo on "So What". Local phrase groups with motivic improvisation are indicated as white rectangles, long-term correlations are indicated by a white ellipse.

similarity values between adjacent phrases. For instance, phrase 12, 13, and 14 (indicated by a white rectangle) form a triple **AA'A"** of varied repetition. As can be seen in Figure 4, these phrases start with the same (sub-)motif but have different endings. Notably, Coltrane employs this creative device quite frequently in his modal solos. There are four motific triples, from which two sets (indicated by a white ellipse in Figure 4 are furthermore related to each other. Long-range correlations like this is one possible way to provide coherence and internal "logic" to an improvisation.

3.3 Histogram distributions

Examining the distributions of parameters such as pitch, interval, inter-onset interval, duration etc. allows direct comparison of solos, players and corpora, and provides hints to meaningful statistical features for classification. For example, in Figure 5, histograms of diatonic chordal pitch classes (classification of pitches according to their relation to the root of the sounding chord) are shown for different solos in our database. It can be observed, for instance, that Louis Armstrong frequently uses thirds in his solo on "West End Blues"[8], which occurs less often with the other players. Presumably, the use of thirds might serve as a feature to discriminate older styles (Swing, New Orleans) from later styles (Bebop, Hardbop, Postbop).

4 Conclusions

The JAZZOMAT project is only at it beginnings right now, but already shows the possibilities of a high-quality database in combination with modern, state-of-art computa-

[8] Rec. 1928, Giants of Jazz CD53001

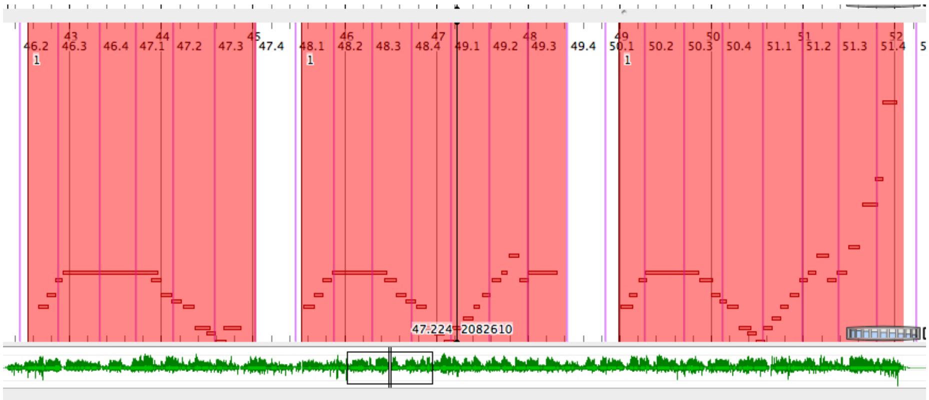

Fig. 4. Sequence of (similar) phrases 12–14 (compare Figure 3) in John Coltrane's solo on "So What".

tional tools. Therefore, we hope to gain new, interesting and statistically well-founded insights in the internal mechanisms of the jazz improviser's mind.

5 Acknowledgments

The JAZZOMAT project is supported by a grant DFG-PF 669/7-1 ("Melodisch-rhythmische Gestaltung von Jazzimprovisationen. Rechnerbasierte Musikanalyse einstimmiger Jazzsoli") by the Deutsche Forschungsgemeinschaft. The authors would like to thank all jazz and musicology students participating in the transcription and annotation process.

References

1. Martin Pfleiderer and Klaus Frieler, "The jazzomat project. Issues and methods for the automatic analysis of jazz improvisations," in *Concepts, Experiments, and Fieldwork: Studies in Systematic Musicology and Ethnomusicology*, Rolf Bader, Christiane Neuhaus, and Ulrich Morgenstern, Eds., pp. 279–295. Peter Lang, Frankfurt/M., Bern, 2010.
2. George Tzanetakis, Ajay Kapur, W. Andrew Schloss, and Matthew Wright, "Computational ethnomusicology," *Journal of Interdisciplinary Music Studies*, vol. 1, no. 2, pp. 1–24, 2007.
3. David Huron, *Sweet Anticipation. Music and the Psychology of Expectation*, MIT Press, Cambridge, 2006.
4. Marcus. T. Pearce, Daniel Müllensiefen, and Geraint. A. Wiggins, "A comparison of statistical and rule-based models of melodic segmentation," in *Proceedings of the Ninth International Conference on Music Information Retrieval*, Philadelphia, USA, 2008, pp. 89–94, Drexel University.
5. Daniel Müllensiefen, Geraint Wiggins, and David Lewis, "High-level feature descriptors and corpus-based musicology: Techniques for modelling music cognition," in *Hamburger Jahrbuch für Musikwissenschaft: Vol. 25. Systematic and Comparative Musicology*, Albrecht Schneider, Ed., Frankfurt/M., Bern, 2008, pp. 133–156, P. Lang.

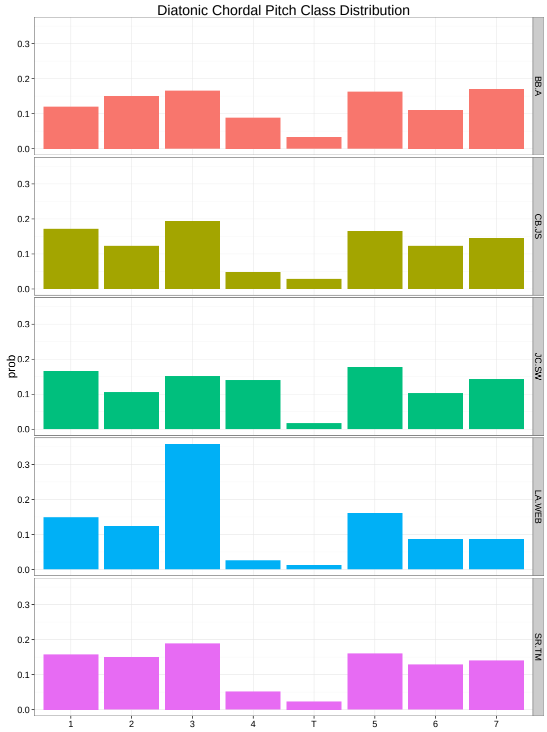

Fig. 5. Diatonic chordal pitch class histograms for a sample of different soli. Pitches are classified here according to their relation to the root of the sounding chord. 1 denotes the root, 2 denotes major and minor seconds, 3 minor and major third etc., T is the tritone interval. BB.A = Bob Berg "Angles" (Rec. 1993, Stretch Records STD-1105), CB.JS = Clifford Brown "Joy Spring" (Rec. 1945, Nippon Phonogram Co., Ldt0 838 308-2), JC.SW = John Coltrane "So What" (Rec. 1959, CBS 460603 2), LA.WEB = Louis Armstrong "West End Blues" (Rec. 1928, Giants of Jazz CD53001), SR.TM = Sonny Rollings "Tenor Madness" (Rec. 1956, CD Company Italy CD53061)

6. Darrell Conklin and Christina Anagnostopoulou, "Segmental pattern discovery in music," *Informs. Journal of Computing*, vol. 18, pp. 285–293, 2006.
7. Olivier Lartillot and Petri Toiviainen, "Motivic matching strategies for automated pattern extraction," *Musicae Scientiae*, vol. 11, no. 1 Suppl, pp. 281–314, Mar. 2007.
8. David Meredith, Kjell Lemström, and Geraint A. Wiggins, "Algorithms for discovering repeated patterns in multidmensional representations of polyphonic music," *Journal of New Music Research*, vol. 31, no. 4, pp. 321–345, 2002.
9. Paul F. Berliner, *Thinking in Jazz. The Infinite Art of Improvisation.*, University of Chicago Press, Chicago, 1994.
10. Martin Pfleiderer, "Improvisieren – ästhetische Mythen und psychologische Einsichten," in *Improvisieren. 8. Darmstädter Jazzforum 2003*, Wolfram Knauer, Ed., Hofheim, 2004, pp. 81–99, Wolke.
11. Thomas Owens, *Charlie Parker. Techniques of Improvisation*, Ph.D. thesis, University of California, Los Angeles, 1974.

12. Pierre-Yves Rolland, "Discovering patterns in musical sequences," *Journal of New Music Research*, vol. 28, no. 4, pp. 334–350, 1999.
13. Cory McKay, "Automatic genre classification of midi recordings," M.S. thesis, McGill University, Canada, 2004.
14. Olivier Lartillot, "Motivic pattern extraction in symbolic domain," in *Intelligent Music Information Systems: Tools and Methodologies*, Jialie Shen, John Shepard, and Bin Cui Ling Liu, Eds., chapter 7, pp. 233–257. IGI Global, 2007.
15. Vladimir Karbusicky, *Kosmos-Mensch-Musik. Strukturalistische Anthropologie des Musikalischen*, Verlag R. Kramer, 1990.