WHO'S PLAYING THAT SOLO? RECOGNIZING JAZZ MUSICIANS BY THEIR "UNIQUE SOUND"

Jakob Abeßer[1,2], Martin Pfleiderer[1], Klaus Frieler[1], Estefanía Cano[2]

[1] Jazzomat Research Project, University of Music "Franz Liszt" Weimar, Germany
[2] Semantic Music Technologies Group, Fraunhofer IDMT, Ilmenau, Germany

24 September 2016
The Jazzomat Research Project: Perspectives for Computational Jazz Studies

Outline

Motivation

”Jazz musicians have always acknowledged the importance of developing a ***unique*** *stylistic* ***voice*** *as a way of transcending from imitation and assimilation into innovation.”*

(Berliner, 1994)

Motivation

"Jazz musicians have always acknowledged the importance of developing a ***unique*** *stylistic* ***voice*** *as a way of transcending from imitation and assimilation into innovation."*

(Berliner, 1994)

- What make jazz soloists recognizable by listeners?
  - Syntactical features (pitch, interval, rhythm, harmonic / metrical context)
  - Non-syntactical / expressive features (micro-timing, dynamics, intonation, pitch modulation, **timbre / sound**)

Motivation

- Which factors affect the recognition?[1]
  - **Instrument sound of performer**
  - Expressive tone modifications
    - Dynamics, articulations, pitch modulations, microtiming, ...
  - Sound of accompanying instruments (rhythm section)
  - Recording conditions
    - Recording setup, microphone characteristics, recording year, ...

[1] Audio example

Motivation

- Goals of this study
  - Focus on instrument sound as complementary part to other expressive features for jazz performance analysis
  - Focus on trumpet (tp), alto saxophone (as), and tenor saxophone (ts)
  - Evaluation scenarios: artist classification & similarity

Motivation

- Application Scenarios
  - Verification of known & discovery of unknown timbral similarities between jazz soloists
  - Automatic performer identification in jazz recordings
    - Content-based metadata enrichment & cleanup for jazz archives
  - Evolution of artist sound
    - Imitation strategies among jazz soloists

Outline

Related Work

- Timbre Research
  - Timbre = Difference between sounds of the same pitch & loudness
  - Among ”spectrographic attributes” [Benadon, 2003] with microtemporal deviations (microtiming) & expressive nuances (vibrato & pitch bending)
  - Previous studies often use perceptual scaling to indentify *underlying acoustic dimensions*
    - Note envelope, temporal change in spectrum
    - Rise time & quality of note attack
    - Spectral centroid

Related Work

- Jazz Performer Identification
  - [Benadon, 2003]
    - Music performance = *Calligraphic* (pitch, rhythm, contour) vs. *spectrographic* (timbre, microtiming, pitch modulation) aspects
    - Tenor saxophonist identification (*John Coltrane, Dexter Gordon, Sonny Rollins, Wayne Shorter*)
    - Listening tests (16 short 2-5 note sequences from post-bebop recordings)
    - Results: 7 subjects show average recognition score of 11.9/16
  - [Ramirez et al., 2010]
    - Analysis of monophonic saxophone solo recordings (4 standards, 2 tempi, 3 performers)
    - Spectral tone model, inter-note and intra-note features
    - Deviation patterns from pitch, timing, amplitude, timbre
    - Up to 98 % classification accuracy

Related Work

- Jazz Performer Identification
  - [Lazar and Lesk, 2016b]
    - Comparison of 3 trumpet performances of *St. Louis Blues* by *Louis Amstrong, Harry James, and Wynton Marsalis*)
    - Spectrogram-based observation of characteristic timbre propoerties
    - Note-level: sound fuzziness & sound clarity
    - Segment-level: sound complexity & note envelopes (rising time)
  - [Abeßer et al., 2015, Lazar and Lesk, 2016a]
    - Importance on vibrato features for jazz performer classification
- Similar approaches for identification of singers, cellists, piano players

Outline

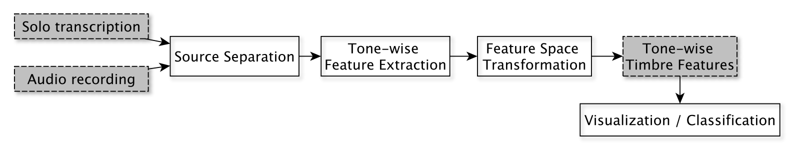
1. Motivation
2. Related Work
3. Proposed Method
4. Experiments & Results
    - Dataset
    - Artist Classification
    - Artist Sound Similarity
5. Conclusion & Outlook

Proposed Method

- How to analyze the ”sound” of a jazz soloist?
  - Data source: solos in commercial jazz recordings (multiinstrumental, polyphonic)
  - Use high-quality score information (Weimar Jazz Database) for “tone localization”
  - Isolate solo instrument from mixture signal (source separation)
  - Tone-level spectrogram-based analysis
  - Quantify different spectral & temporal properties of tones (state-of-the-art timbre features)
  - Use machine learning methods to learn artist-specific timbre patterns

Proposed Method

- Processing Flowchart

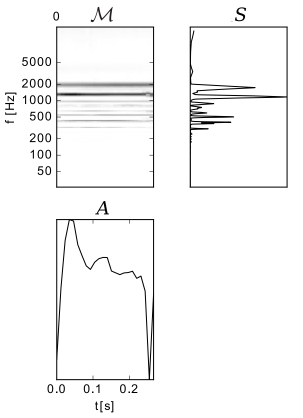


Proposed Method

- Pitch-informed solo & accompaniment separation [Cano et al., 2014]
  - Goal → isolate improvising solo instrument from accompaniment instruments (rhythm section)
  - Iterative **spectral modeling** of the solo instrument in the spectal domain
  - Includes musical instrument characteristics such as **common amplitude modulation**, **inharmonicity** & **magnitude and frequency smoothness**

Proposed Method

- Feature Extraction
  - One-Component Non-Negative Matrix Factorization NMF
    - Represent tone spectrogram M as product of spectral envelope S and temporal activation A



Proposed Method

- Feature Extraction
  - Spectral Features
    - Representation: S
    - $\rightarrow$ Statistics (centroid, spread, skewness, kurtosis)
    - $\rightarrow$ Shape (decrease, slope, flatness, roll-off)
    - Mel-Frequency Cepstral Coefficients (MFCC)
    - Spectral Contrast (octave-based, shape-based)
    - Inharmonicity
    - Relative Harmonic Magnitudes
    - Odd-to-even ratio
    - Tristimulus 1-3

Proposed Method

- Feature Extraction
  - Temporal Features
    - Representations: A, ΔA, 10-bin histogram over ΔA
    - $\rightarrow$ Statistics (centroid, spread, skewness, kurtosis)
    - $\rightarrow$ Shape (decrease, slope, flatness, roll-off)
    - Relative attack part length
    - Log attack / decay time
    - Multi-resolution gamma filterbank (correlation with prototype envelopes)

Outline

1. Motivation
2. Related Work
3. Proposed Method
4. **Experiments**
    - Dataset
    - Artist Classification
    - Artist Sound Similarity
5. Conclusion & Outlook

Experiments

- Dataset
  - Solos (audio + manual solo transcriptions) taken from *Weimar Jazz Database*[2]
  - Performer selection $\rightarrow \geq 4$ solos / performer
    - tp (7) $\rightarrow$ *Chet Baker, Dizzy Gillespie, Freddie Hubbard, Kenny Dorham, Miles Davis, Roy Eldridge*
    - ts (13) $\rightarrow$ *Bob Berg, Coleman Hawkins, David Murray, Dexter Gordon, Don Byas, John Coltrane, Joe Henderson, Joshua Redman, Lester Young, Michael Brecker, Sonny Rollins, Stan Getz, Wayne Shorter*
    - as (8) $\rightarrow$ *Art Pepper, Benny Carter, Cannonball Adderley, Charlie Parker, Lee Konitz, Ornette Coleman, Paul Desmond, Steve Coleman*
  - Note selection $\rightarrow$ first 100 tones / solo with ≥ 100 ms duration

[2] http://jazzomat.hfm-weimar.de/dbformat/dboverview.html

Artist classification

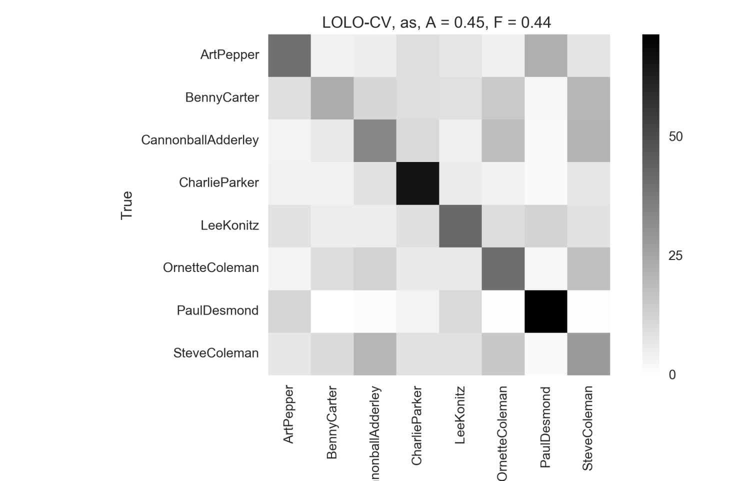
- Cross-validation strategy
  - Leave-one-label-out cross-validation (LOLO-CV)
    - Split of tones to training & test set strictly based on solo id
    - Realistic assumption: some solos of artist are known
- Optional majority voting over all tones of a given solo

Artist classification

Instrument	tp	as	ts
Number of Performers	7	8	13
Random baseline	0.14	0.13	0.08
LOLO-CV	0.36	0.45	0.28
LOLO-CV (majority voting)	**0.66**	**0.78**	**0.51**

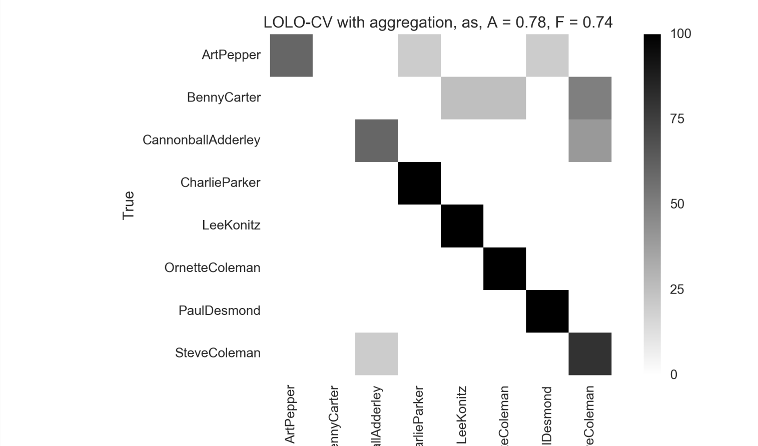
Artist classification (as)

- Example: alto saxophone - tone-wise classification



Artist classification (as)

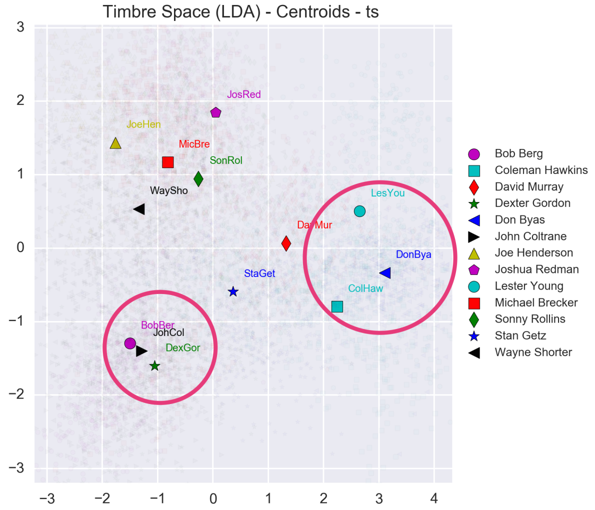
- Example: alto saxophone - solo-wise classification (majority voting over tones)



Artist Sound Similarity

- Explore reduced version of timbre feature space
- Transform feature space (127 dim. $\rightarrow$ 2 dim.) using Linear Discriminant Analysis (LDA) by
  - maximizing the variances between performers
  - minimizing the variances within performers
  - (Feature dimensions are linear combinations of existing features)
- Interpretation
  - Distance between class centroids $\rightarrow$ timbre similarity between performers
  - Clusters $\rightarrow$ Performers with similar sound

Artist Sound Similarity (ts)

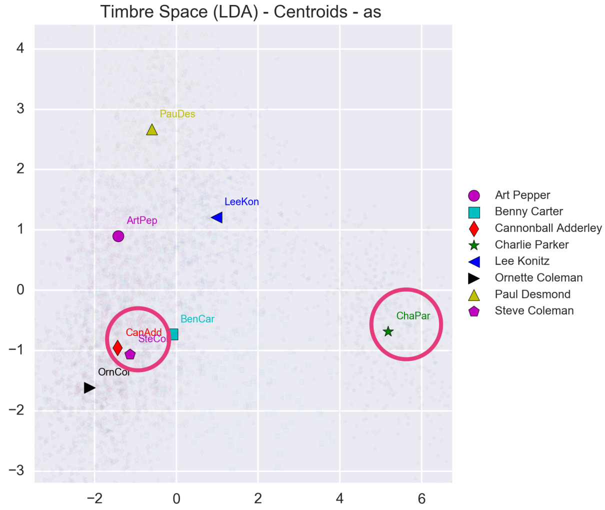


Artist Sound Similarity (ts)

- Obervations
  - Group 1: Coleman Hawkins, Lester Young, Don Byas
    - Swing tenor sax players
  - Cluster 2: John Coltrane, Bob Berg[3], Dexter Gordon
    - Dexter Gordon = early influence of John Coltrane[4]
    - Bob Berg: "The example of John Coltrane was a major influence on his playing [...]."[5]

[3] Audio example
[4] C. Woideck: John Coltrane: Development of a Tenor Saxophonist, 1950–1954, Jazz Perspectives, Vol. 2, Issue 2, 2008
[5] http://www.jazzhouse.org/gone/lastpost2.php3?edit=1039185731

Artist Sound Similarity (AS)

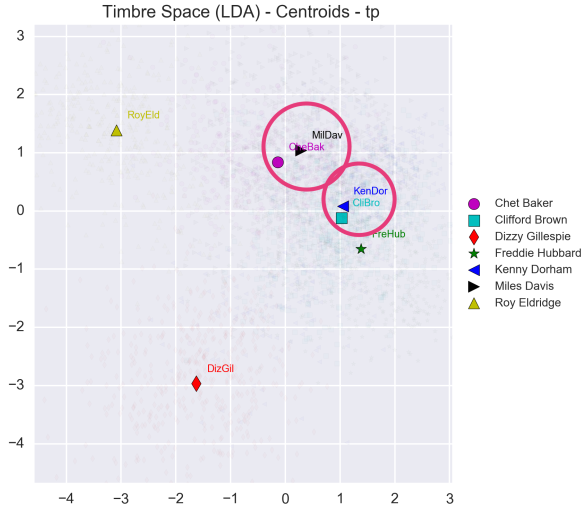


Artist Sound Similarity (AS)

- Obervations
  - Cluster: Steve Coleman, Cannonball Adderley[6]
  - Unique sound: Charlie Parker

[6] Audio example

Artist Sound Similarity (tp)



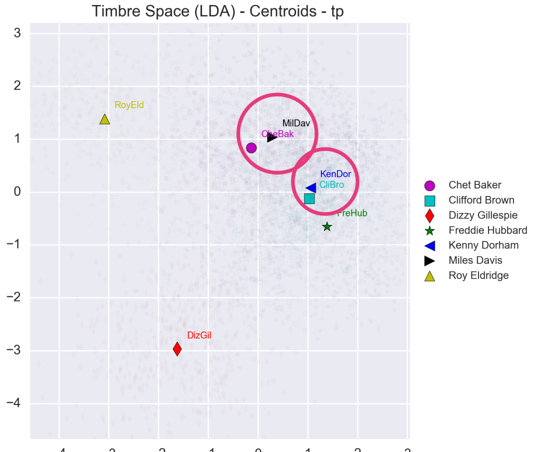
Artist Sound Similarity (tp)

- Observations
  - Unique sound: Dizzy Gillespie & Roy Eldridge
  - Cluster 1: Chet Baker, Miles Davis[7]
    - “[..] his small tone and limited range will remind some listeners of Miles Davis.”[8]
  - Cluster 2: Kenny Dorham, Clifford Brown[9]
    - Bebop trumpet players

[7] Audio example
[8] http://jazztimes.com/articles/20336-chet-baker-his-life-and-music-jeroen-de-valk
[9] Audio example

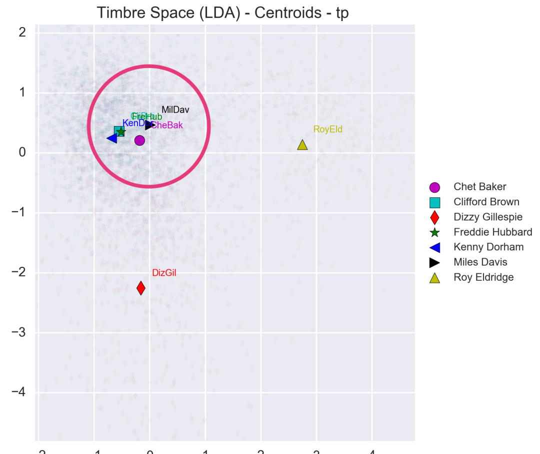
- Outlook: Influence of seperation method
  - More “strict” separation can lead to losing valuable timbre information (noise / transient properties)
  - More “loose” separation $\rightarrow$ Classifier might learn sound of recording / accompanying instruments
  - Example: Separability of trumpet players

- Spectral Harmonic Filtering



Artist Sound Similarity (tp)

- Source Separation



Outline

Conclusion

- New approach for score-informed timbre analysis in jazz solos
- Weimar Jazz Database allows for “large-scale” evaluation experiments
- Automatic soloist classification improved by
  - Knowledge of solo instrument
  - Majority voting over solos
- Timbre feature space confirms known / reveals unknown similarities between jazz soloists

Outlook

- Increase dataset size (artists, solos)
- Automatic Feature Learning (DNN)
- Quantify importance of different factors for artist recognition
  - Harmonic components (fundamental frequency, partials)
  - Noise components (attack transients, noise, subharmonic components)
- Combine timbre & style features (audio & symbolic) for artist & style classification
- Listening test $\rightarrow$ human performance?
- Artist sound evolution $\rightarrow$ imitation strategies among soloist

Bibliography I

Abeßer, J., Cano, E., Frieler, K., Pfleiderer, M., and Zaddach, W.-G. (2015).
Score-informed analysis of intonation and pitch modulation in jazz solos.
In *Proceedings of the 16th International Society for Music Information Retrieval Conference (ISMIR)*, pages 823–829, Málaga, Spain.

Benadon, F. (2003).
Spectrographic and calligraphic cues in the identification of jazz saxophonists.
In Kopiez, R., Lehmann, A. C., Wolther, I., and Wolf, C., editors, *Proceedings of the 5th Triennial ESCOM Conference*, pages 246–249, Hanover, Germany.

Cano, E., Schuller, G., and Dittmar, C. (2014).
Pitch-informed solo and accompaniment separation: towards its use in music education applications.
EURASIP Journal on Advances in Signal Processing, pages 1–19.

Lazar, J. G. and Lesk, M. (2016a).
It must be Louis 'cause Miles don't shake like that: Towards identifying jazz trumpeters by vibrato.
In *Extended Abstracts for the Late-Breaking Demo Session of the 17th International Society for Music Information Retrieval Conference*.

Bibliography II

Lazar, J. G. and Lesk, M. (2016b).
Who's playing? towards machine-assisted identification of jazz trumpeters by timbre.
In *Proceedings of the IConference 2016*.

Ramirez, R., Maestre, E., and Serra, X. (2010).
Automatic performer identification in commercial monophonic jazz performances.
Pattern Recognition Letters, 31:1514–1523.

Thank you very much for your attention!

Questions?