JAZZ SOLO ANALYSIS BETWEEN MUSIC INFORMATION RETRIEVAL, MUSIC PSYCHOLOGY, AND JAZZ RESEARCH.

Part III - Score-informed Solo Analysis

Wolf-Georg Zaddach, Klaus Frieler, **Jakob Abeßer**

August, 7, ISMIR 2016

University of Music “Franz Liszt” Weimar, Germany

→ `https://github.com/jazzomat/knowhow` (Tutorial slides, references, web resources)

Outline

- Introduction
- Method
- Dataset
- Experiments
  - Tuning
  - Intonation & Pitch Modulation
  - Dynamics
- Summary & Outlook

Outline

- Introduction
- Method
- Dataset
- Experiments
  - Tuning
  - Intonation & Pitch Modulation
  - Dynamics
- Summary & Outlook

Introduction

- Audio signal processing / MIR methods can **assist/accelerate/stimulate** musicological research (amount of data, processing time)
- Results of automated analysis procedures often **require expert knowledge** for thorough interpretation
- Audio-based analysis in the **Jazzomat Research Project**
  - Get insights into **stylistic characteristics** of improvising jazz musicians
  - Focus on **non-syntactic / expressive properties** of music performance (intonation, pitch modulation, dynamics, timbre, micro-timing)

Outline

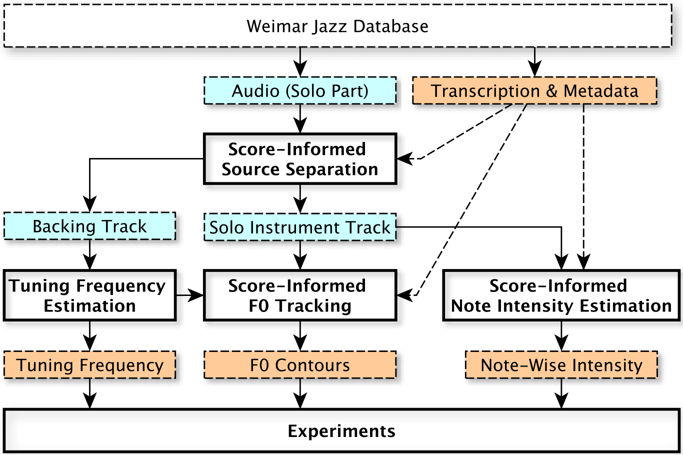
- Introduction
- Method
- Dataset
- Experiments
    - Tuning
    - Intonation & Pitch Modulation
    - Dynamics
- Summary & Outlook

Method

- High quality **solo transcriptions** (Weimar Jazz Database)
- Pitch-informed **solo & accompaniment separation** (solo & backing track)
- **Tuning** estimation (backing track)
- Note-wise tracking of **fundamental frequency and intensity contours** (solo track)
- **Statistical analyses** (informed by symbolic analysis results & solo metadata)

Method

Framework overview



Method

- Pitch-informed solo & accompaniment separation [Cano et al., 2014]
  - Goal $\rightarrow$ 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**

Method

- Pitch-informed solo & accompaniment separation [Cano et al., 2014]
  - Examples
    - Bob Berg - Angles (ts, Postbop, 1993, 270.3 bpm)
    - Clifford Brown - Joy Spring (tp, Hardbop, 1945, 161.2 bpm)
    - Lester Young - Body And Soul (ts, Swing, 1942, 78.6 bpm)

Outline

- Introduction
- Method
- **Dataset**
- Experiments
  - Tuning
  - Intonation & Pitch Modulation
  - Dynamics
- Summary & Outlook

Dataset

- Selection from Weimar Jazz Database
  - Reed and brass solo instruments (tp, cl, as, ts, ss, tb, cor)
  - Artists with ≥ 3 solos
  - Note duration ≥ 50 ms
  - **104,964 note events, 264 solos, 47 performers**
- Metadata
  - Composition info $\rightarrow$ artist, title, instrument, recording date
  - Note annotations $\rightarrow$ onset, duration, pitch, salient pitch modulations (vibrato, pitch bends, etc.)
  - Contextual annotations $\rightarrow$ metrical position, musical phrases

Dataset

Performer	Instrument	# Solos	# Notes
Art Pepper	as, cl	6	3482
Ben Webster	ts	5	852
Benny Carter	as	5	1750
Benny Goodman	cl	7	1966
Bix Beiderbecke	cor	4	518
Bob Berg	ts	6	4000
Buck Clayton	tp	3	561
Cannonball Adderley	as	5	2475
Charlie Parker	as	6	1606
Chet Baker	tp	6	1079
Clifford Brown	tp	7	2890
Coleman Hawkins	ts	6	2465
David Liebman	ss, ts	5	3210
David Murray	ts	6	2810
Dexter Gordon	ts	5	3702
Dickie Wells	tb	3	387
Dizzy Gillespie	tp	5	1384
Don Byas	ts	7	1928
Eric Dolphy	as	3	1500
Fats Navarro	tp	4	876
Freddie Hubbard	tp	6	2016
Gerry Mulligan	bs	3	1049
Hank Mobley	ts	3	1462
J.J. Johnson	tb	5	2215
Joe Henderson	ts	6	3534
Joe Lovano	ss, ts, ts-c	6	4116
John Coltrane	ss, ts	11	8042
Joshua Redman	ts	5	2344
Kenny Dorham	tp	6	1922
Kid Ory	tb	3	174
Lee Konitz	as	5	2202
Lester Young	ts	6	1452
Louis Armstrong	cor, tp	6	782
Michael Brecker	ts	6	4076
Miles Davis	tp	8	2377
Ornette Coleman	as	5	2718
Paul Desmond	as	8	2119
Roy Eldridge	tp	6	1643
Sidney Bechet	ss	3	695
Sonny Rollins	ts	12	4797
Sonny Stitt	as, ts	4	1239
Stan Getz	ts	6	3129
Steve Coleman	as	7	2776
Steve Lacy	ss	5	1661
Steve Turre	tb	3	1038
Wayne Shorter	ts	10	3510
Woody Shaw	cor, tp	6	2435
Total		**264**	**104964**

(a) (b)

Figure: Dataset overview

Outline

Tuning - Related Work

- Tuning
  - **Adjustment of pitch frequencies** to given **reference frequency** (often 440 Hz)
  - Allows for **coordinating** the **intonation of multiple instruments** during a performance
  - **Deviations of jazz recordings** from standard tuning possible due to
    - No authoritative international standard before 1955
    - Tuning deviations of pianos
    - Speed variations of gramophones or tape recorders (up to the 1940s)

Tuning - Related Work

- Tuning Estimation
  - Important **pre-processing step** for MIR tasks (key and chord estimation, music transcription)
  - Approaches
    - High-resolution **interval histograms based on peak frequencies** [Serrá et al., 2011]
    - Match peak frequencies to **tuning/temperament templates** [Dixon et al., 2011]
    - Use **adjustable semitone filterbanks** to evaluate different tuning hypotheses [Lerch, 2006, Müller and Ewert, 2011]
    - Analysis of **phase spectrogram** [Mauch, 2010]
    - Analysis of **tuning deviations in the complex plane** using **circular statistics** [Dressler and Streich, 2007]

Tuning - Method

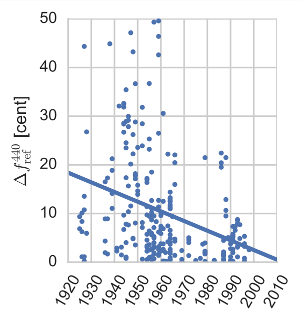
- Tuning Frequency Estimation
  - Compared available implementations by
    - Müller and Ewert [Müller and Ewert, 2011] (*Chroma Toolbox*)[1]
    - Mauch [Mauch, 2010] (*NNLS Chroma Vamp Plugin*)[2]
  - **No groundtruth**
  - **High agreement** between methods
    - Sample correlation of $r = 0.96$ ($p < .001$)
    - RMSE = 0.13 cent
    - 7 outlier ($\Delta f > 50$ cent)
  - Used NNLS plugin (faster)

[1] http://resources.mpi-inf.mpg.de/MIR/chromatoolbox/
[2] http://www.isophonics.net/nnls-chroma

- Tuning Deviations by Recording Year
  - Hypothesis
    - **Tuning deviations** of jazz recordings from 440 Hz are **stronger in the first half of the 20th century**
  - Approach
    - Tuning frequency estimation based on backing track (rhythm section)
    - **Deviation from ideal tuning frequency**

$$\Delta f_{\mathrm{ref}}^{440} = \left| 1200 \log_2 \frac{f_{\mathrm{ref}}}{440} \right|$$

- Tuning Deviations by Recording Year
  - Results
    - **Negative correlation** between $\Delta f_{\mathrm{ref}}^{440}$ and the recording year : $r = -0.33$ ($p < 0.001$).
    - "**Turning point**" around 1960 $\rightarrow$ (slightly delayed) adoption of the standardized 440 Hz tuning frequency issued by the International Standards Organization in 1955 (?)



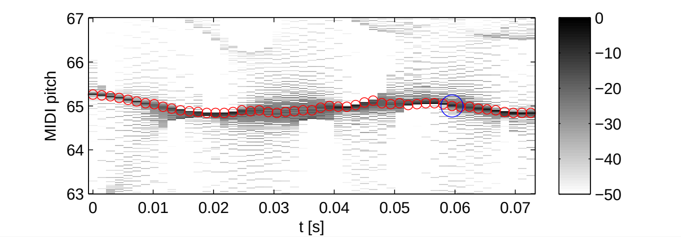
Pitch Modulation & Intonation - Related Work

- Notes are rarely played/sung with constant pitch
- **Pitch modulation** techniques are an important part of expressive music performance (jazz $\rightarrow$ fall-off, slide up/down, vibrato) [Berliner, 1994]
- Analysis requires estimation of the tone-wise **fundamental frequency** (f_0) **contours**
- Problem $\rightarrow$ interference between music sources
- Methods
  - f_0 contour formation as part of pitch estimation in automatic melody transcription algorithms [Salamon et al., 2012, Dressler, 2011]
  - Score-informed: prior score information as guideline to restrict the search space for the f_0 contour [Abeßer et al., 2014]

Pitch Modulation & Intonation - Related Work

- Vibrato → periodic pitch change around target pitch (vocalists [Sundberg, 1994], string and wind instrument players [Liley, 2007])
- Jazz → vibrato shaping (speed, amount, intra-note start time) [Pfleiderer, 2009]
- Methods for detection/analysis of vibrato
  - Pitch contour segmentation & heuristics [Miryala et al., 2013]
  - Contour-features & classification [Hsu and Jang, 2010, Abeßer et al., 2010, Salamon et al., 2012, Chen et al., 2015]
  - Template-based [Driedger et al., 2016] → vibrato salience spectrogram

- Spectral Estimation
  - STFT magnitude spectrogram ($b = 2048$, $h = 128$, $z = 8$, $f_s = 44.1$ kHz)
  - Log-spaced frequency axis ($\pm$ 2 semitones around target pitch, 25 bins per semitone)
  - **Magnitude reassignment** towards **instantaneous frequency** $\hat{f}$ [Abe et al., 1995]
- Contour tracking
  - Frame-wise **forwards-backwards tracking** of f_0 contour (**peak detection** & **proximity**)



Pitch Modulation & Intonation - Method

- Method
  - Example (Sonic Visualiser)
    - Ben Webster - Night And Day (ts, Swing, 1956, 181.8 bpm)

Pitch Modulation & Intonation - Experiments

- Artist-Specific Pitch Modulation
  - Hypothesis
    - Important aspects of “personal sound” of jazz musicians [Berliner, 1994]
    - Statistics over large number of tones reveal **artist-specific characteristics**
  - Method
    - **Modulation range** (measure of pitch stability) $\rightarrow f_0$ variation per tone, average interquartile range (IRQ) over f_0 contours in cent

Pitch Modulation & Intonation - Experiments

- Artist-Specific Pitch Modulation
  - Results

#	Performer	Instrument	Average IQR [cent]
1	Steve Lacy	ss	17.4
2	Benny Goodman	cl	17.5
3	Woody Shaw	cor, tp	19.8
4	Dizzy Gillespie	tp	21.2
5	Freddie Hubbard	tp	21.5
6	Art Pepper	as, cl	22.1
7	David Liebman	ss, ts	22.1
41	Sonny Rollins	ts	30.4
42	Michael Brecker	ts	30.5
43	Ben Webster	ts	30.8
44	Dickie Wells	tb	30.9
45	Coleman Hawkins	ts	31.4
46	Lester Young	ts	31.7
47	Don Byas	ts	33.7

Pitch Modulation & Intonation - Experiments

- Artist-Specific Intonation
  - Method
    - **Pitch intonation deviation** Δf_0 (pitch deviation) $\rightarrow$ median f_0 deviation from ground-truth pitch frequency (global tuning frequency taken into account)
    - **Intonation class**: “flat” (-): $\Delta f_0 < -25$ cent, “sharp” (+): $\Delta f_0 > 25$ cent, “normal” (o)
    - **Intonation tendency**: $T = (N_{\text{sharp}} - N_{\text{flat}})/(N_{\text{sharp}} + N_{\text{flat}})$
    - **Intonation tendency class**: “flat” (-): $T < -0.15$, “sharp” (+): $T > 15$, “no tendency” (o)
    - **Intonation-goodness-measure**: $T = N_{\text{normal}}/N$

Pitch Modulation & Intonation - Experiments

- Artist-Specific Intonation
  - Results

#	Performer	Instrument	N_{on}/N	Intonation Tendency T	Class
1	Benny Goodman	cl	0.81	0.24	+
2	Chet Baker	tp	0.81	0.26	+
3	Bix Beiderbecke	cor	0.80	0.24	+
4	Louis Armstrong	cor, tp	0.79	-0.01	o
5	Freddie Hubbard	tp	0.79	-0.26	-
6	Woody Shaw	cor, tp	0.79	-0.07	o
7	Paul Desmond	as	0.79	0.39	+
41	Ornette Coleman	as	0.68	-0.18	-
42	Eric Dolphy	as	0.68	0.01	o
43	Fats Navarro	tp	0.67	0.18	+
44	Roy Eldridge	tp	0.67	0.56	+
45	Michael Brecker	ts	0.65	0.05	o
46	Don Byas	ts	0.63	0.43	+
47	Charlie Parker	as	0.59	0.53	+

Pitch Modulation & Intonation - Experiments

- Artist-Specific Vibrato Modulation Frequency
  - Hypothesis
    - Modulation frequency (vibrato) is mostly an idiosyncratic part of the personal style of jazz musicians
  - Method
    - Given: **tone-wise f_0-contours** & **human annotations** (vib)
    - Estimation of $f_{\text{mod}} \in [3, 10]$ Hz (Modulation range [5,8] Hz [Fletcher, 2010] + margin) using **autocorrelation** & **peak search**
    - Selection of artists with ≥ 20 vibrato tones
    - Distribution analysis (box plots)

- Artist-Specific Vibrato Modulation Frequency
  - Results
    - Median values in [4, 7] Hz
    - 3 trumpet players with considerable vibrato (L. Amstrong, K. Dorham, R. Eldridge)

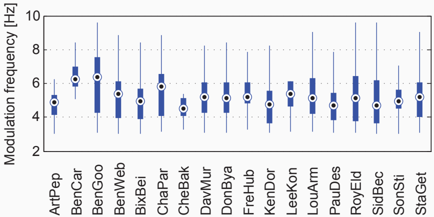


Figure: [Abeßer et al., 2015]

Dynamics - Related Work

- Important aspect of any **musical performance** [Gabrielsson, 1999, Langner and Goebel, 2003]
- **Varying intensity/accentuation** depend on
  - **Structural position in musical phrases** ("local stresses" or "phenomenal accents" [Lerdahl and Jackendoff, 1983])
  - **Metrical position** (syncopation, down-beat vs. off-beat, cross-rhythmic superpositions)
- Precise estimation of single tone intensities in ensemble recordings is complicated
  - Interference/overlap between instrument signals
  - Room/instrument acoustics

Dynamics - Related Work

- Expressive performance analysis [Widmer and Goebl, 2004] focus on on isolated tracks
  - Piano
    - **Score-informed** spectrogram modeling [Ewert and Müller, 2011]
    - **Unsupervised feature learning** for modeling note intensities [Grachten et al., 2014]
    - **Prediction of dynamics** using neural networks [Van Herwaarden et al., 2014]
  - Saxophone
    - **Note amplitude estimation** via spectral modeling synthesis (SMS) [Arcos et al., 1998]
    - **Intra-note segmentation** based on energy envelope contour [Ramirez et al., 2007]

Dynamics - Method

- Score-Informed note intensity estimation
  - Intensity computation [Painter and Spanias, 2000]
    - **Band-wise intensity values** for 24 critical bands

$$I_b(n) = \frac{1}{b} \sum_{k \in [k_{\min,b}, k_{\max,b}]} |X(k,n)|^2 \,. \tag{1}$$

    - **Frame-wise intensity values**

$$I(n) = 90.302 + 10 \log_{10} \sum_{b=1}^{24} I_b(n) \tag{2}$$

    - **Aggregation** over tones (max, median, standard deviation, relative peak position, temporal centroid)
    - **Normalization** of tone intensity values by mapping the 5%–95% percentiles to the interval $[0, 1]$ for each solo

Dynamics - Experiments

- Phrase-Wise Intensity and Pitch Contours
  - **Melodic phrases** contribute to
    - emotional impression
    - melodic perception and memory [Dowling and Fujitani, 1971]
    - melodic classification and similarity judgments [Müllensiefen and Wiggins, 2011]
  - Approach
    - **Contour type classification** (horizontal, ascending, descending, concave, convex)
    - Adaptation of contour-classes by Huron [Huron, 1996] to model longer phrase lengths
    - **Phrase segmentation** $\rightarrow$ Median intensity & pitch values of first 25 %, central 50 %, and final 25 %

Dynamics - Experiments

- - Segment-wise (pitch/intensity) differences

$$\Delta v_i = \begin{cases} \mathrm{sign}(v_{i+1} - v_i), & \text{if } |v_{i+1} - v_i| \geq \Delta v_{\min} \\ 0, & \text{otherwise} \end{cases} \tag{3}$$

  using $\Delta v_{\min} = 0.1(\max_i v_i - \min_i v_i)$.
  - Heuristic

Contour Type	$(\Delta v_1, \Delta v_2)$
Horizontal	$(0, 0)$
Convex	$(1, -1)$
Concave	$(-1, 1)$
Ascending	$(0, 1), (1, 0), (1, 1)$
Descending	$(0, -1), (-1, 0), (-1, -1)$

- Results
  - **Occurance rate** (%) of **pitch and intensity contours** in jazz phrases

Contour Type	Intensity	Pitch
$<$ 4 notes	17.1	17.1
Horizontal	6.0	2.9
Convex	12.9	17.0
Concave	14.3	10.2
Ascending	14.1	17.6
Descending	**35.7**	**35.2**

- Results
  - **Co-occurance rate** (%) of **pitch and intensity contours** in jazz phrases

Pitch Contour Type	Intensity Contour Type				
	Horizontal	Convex	Concave	Ascending	Descending
Horizontal	**14.1**	7.3	15.7	14.1	**48.7**
Convex	6.2	**26.8**	13.5	17.2	**36.4**
Concave	6.9	9.3	**36.6**	13.2	**34.0**
Ascending	7.8	17.4	17.0	**29.0**	**28.8**
Descending	7.0	11.7	13.6	12.1	**55.6**

- Alternating Eighth-Note Accentuations
  - Hypothesis
    - **Accentuation of second eighth-notes** is commonly used to emphasize swing feel
  - Method
    - Identify sequences of ≥ 10 successive eight notes
    - Paired t-test between the intensity values on the second and first eighths (significance)
    - Cohen's d [Cohen, 1988] (effect size)

Dynamics - Experiments

- Alternating Eighth-Note Accentuations
  - Results
    - 33 of 299 solos with **significant intensity differences** for **at least 10** successive eight-note pairs
    - 18 of 33 solos with **higher intensity values for the first eighth notes**
    - **Tendency towards on-beat accentuation**

Dynamics - Experiments

- Alternating Eighth-Note Accentuations
  - Results
    - Solos with highest effect size for **positive intensity differences** (first eights louder than second eights)

Performer	Title	Cohen's d	Significance Level
Benny Goodman	**Tiger Rag**	1.1	***
Coleman Hawkins	Body And Soul	0.9	*
Sidney Bechet	Limehouse Blues	0.9	***
Kenny Dorham	Blues In Be-Bop	0.7	*
Coleman Hawkins	Perdido	0.7	***
Benny Goodman	Whispering	0.7	**
Freddie Hubbard	Speak No Evil	0.6	*
Lee Konitz	Wow	0.6	*
David Liebman	No Greater Love	0.5	*
Woody Shaw	Rosewood	0.5	**

- Alternating Eighth-Note Accentuations
  - Results
    - Solos with highest effect size for **negative intensity differences** (second eights louder than first eights)

Performer	Title	Cohen's d	Significance Level
John Coltrane	Mr. P.C.	-0.2	**
Stan Getz	Blues In The Closet	-0.2	**
Clifford Brown	Daahoud	-0.3	*
Steve Lacy	Easy To Love	-0.3	*
Miles Davis	Airegin	-0.3	*
Miles Davis	Blues By Five	-0.3	*
John Coltrane	Blue Train	-0.4	*
Chet Baker	Long Ago And Far Away	-0.4	*
Kenny Dorham	Punjab	-0.4	*
Steve Turre	**Steve's Blues**	-0.5	*

Dynamics - Experiments

- Examples $\rightarrow$ Sonic Visualiser

Dynamics - Experiments

- Method
  - Example (Sonic Visualiser)
    - Benny Goodman - Tiger Rag (cl, Swing, 1936, 268.5 bpm)
    - Steve Turre - Steve's Blues (tb, Postbop, 1987, 204.5 bpm)

Outline

- Introduction
- Method
- Dataset
- Experiments
  - Tuning
  - Intonation & Pitch Modulation
  - Dynamics
- Summary & Outlook

Summary & Outlook

- Score-informed framework for analyzing jazz solos
- Exploit MIR methods from source separation and automatic music transcription for ensemble recording analysis
- Analysis of large database to investigate musicological hypotheses
- Audio analysis as complementary part to symbolic analysis
- Publication of database, raw analysis data (f_0 and intensity contours, contour features etc.), tools, and Python implementations

Summary & Outlook

- Future research directions
  - Timbre - “Sound” of Jazz Musicians
  - Bass transcription & harmonic analysis
  - Microtiming of Soloist vs. Drummer vs. Bass player
  - Unsupervised search for pitch modulation prototypes (f_0 dips, variable vibrato, glissandi)
  - Jazz performer identification using symbolic & audio features

Summary & Outlook

- Resources
  - Github
    - https://github.com/jazzomat/knowhow (Tutorial slides, references, web resources)
    - https://github.com/jazzomat/article_2016 (Raw analysis data)
  - Pymus Python Package
    - https://pypi.python.org/pypi/pymus/
  - Jazzomat Research Project
    - http://jazzomat.hfm-weimar.de/
  - 2nd International Jazzomat Research Workshop - Perspectives for Computational Jazz Studies (September 20-22, 2016, Weimar, Germany)
    - ../workshop2016/workshop2016.html

Bibliography I

Abe, T., Kobayashi, T., and Imai, S. (1995).
Harmonics tracking and pitch extraction based on instantaneous frequency.
In *Proceedings of the Int. IEEE Conference on Acoustics, Speech and Signal Processing (ICASSP-95)*, pages 756–759, Detroit, USA.

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.

Abeßer, J., Lukashevich, H., and Schuller, G. (2010).
Feature-based extraction of plucking and expression styles of the electric bass guitar.
In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 2290–2293.

Abeßer, J., Pfleiderer, M., Frieler, K., and Zaddach, W.-G. (2014).
Score-informed tracking and contextual analysis of fundamental frequency contours in trumpet and saxophone jazz solos.
In *Proceedings of the 17th International Conference on Digital Audio Effects (DAFx-14) (submitted)*.

Bibliography II

Arcos, J. L., de Mántaras, R. L., and Serra, X. (1998).
Saxex a case-based reasoning system for generating expressive musical performances.
Journal of New Music Research, 27:194–210.

Berliner, P. F. (1994).
Thinking in Jazz: The Infinite Art of Improvisation.
Chicago Studies in Ethnomusicology. University of Chicago Press.

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.

Chen, Y.-P., Su, L., and Yang, Y.-H. (2015).
Electric guitar playing technique detection in real-world recordings based on f0 sequence pattern recognition.
In *Proceedings of the 16th International Society for Music Information Retrieval Conference (ISMIR)*, pages 708–714, Málaga, Spain.

Cohen, J. (1988).
Statistical Power Analysis for the Behavioral Sciences.
New York University.

Bibliography III

Dixon, S., Tidhar, D., and Benetos, E. (2011).
The temperament police: the truth, the ground truth, and nothing but the truth.
In *Proceedings of the 12th International Society for Music Information Retrieval Conference (ISMIR)*, pages 281–286, Miami, USA.

Dowling, W. J. and Fujitani, D. (1971).
Contour, interval, and pitch recognition in memory for melodies.
The Journal of the Acoustical Society of America, 49(2):524+.

Dressler, K. (2011).
Pitch estimation by the pair-wise evaluation of spectral peaks.
In *Proceedings of the 42nd AES International Conference on Semantic Audio*, pages 1–10, Ilmenau, Germany.

Dressler, K. and Streich, S. (2007).
Tuning frequency estimation using circular statistics.
In *Proceedings of the 8th International Conference on Music Information Retrieval (ISMIR)*, pages 357–360, Vienna, Austria.

Bibliography IV

Driedger, J., Balke, S., Ewert, S., and Müller, M. (2016).
Template-based vibrato analysis of music signals.
In *Proceedings of the International Conference on Music Information Retrieval (ISMIR)*, New York, USA.

Ewert, S. and Müller, M. (2011).
Estimating note intensities in music recordings.
In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 385–388, Praha, Czech Republic.

Fletcher, N. H. (2010).
Vibrato in music – physics and psychophysics.
In *Proceedings of the International Symposium on Music Acoustics*, pages 1–4.

Gabrielsson, A. (1999).
Psychology of Music, chapter Music Performance, pages 501–602.
Academic Press, San Diego, second edition edition.

Grachten, M., Cancino Chacón, C. E., and Widmer, G. (2014).
Analysis and prediction of expressive dynamics using Bayesian linear models.
In *Proceedings of the 1st international workshop on computer and robotic Systems for Automatic Music Performance*, pages 545–552.

Bibliography V

Hsu, C.-L. and Jang, J.-S. R. (2010).
Singing pitch extraction by voice vibrato/tremolo estimation and instrument partial deletion.
In *Proceedings of the 11th International Society for Music Information Retrieval Conference (ISMIR)*, pages 525–530, Utrecht, Netherlands.

Huron, D. (1996).
The melodic arch in Western folksongs.
Computing in Musicology, 10:3–23.

Langner, J. and Goebel, W. (2003).
Visualizing expressive performance in tempo–loudness space.
Computer Music Journal, 27(4):69–83.

Lerch, A. (2006).
On the requirement of automatic tuning frequency estimation.
In *Proceedings of the 7th Internations Conference on Music Information Retrieval (ISMIR)*, Victoria, Canada.

Lerdahl, F. and Jackendoff, R. (1983).
A generative theory of tonal music.
The MIT press, Cambridge. MA.

Bibliography VI

Liley, T. (2007).
Vibrato: Enhancing the sound.
Clarinet & Saxophone Magazine.

Mauch, M. (2010).
Automatic Chord Transcription from Audio Using Computational Models of Musical Context.
PhD thesis, Queen Mary University of London, London.

Miryala, S. S., Bali, K., Bhagwan, R., and Choudhury, M. (2013).
Automatically identifying vocal expressions for music transcription.
In *Proceedings of the 14th International Society for Music Information Retrieval Conference (ISMIR)*, pages 239–244, Curitiba, Brazil.

Müllensiefen, D. and Wiggins, G. (2011).
Polynomial functions as a representation of melodic phrase contour.
In Schneider, A. and von Ruschkowski, A., editors, *Systematic Musicology: Empirical and Theoretcial Studies*, pages 63–88. Peter Lang, Frankfurt a.M.

Müller, M. and Ewert, S. (2011).
Chroma toolbox: Matlab implementations for extracting variants of chroma-based audio features.
In *Proceedings of the 12th International Society for Music Information Retrieval Conference (ISMIR)*, pages 215–220, Miami, USA.

Painter, T. and Spanias, A. (2000).
Perceptual Coding of Digital Audio.
Proceedings of the IEEE, 88(4):451–515.

Pfleiderer, M. (2009).
Five Perspectives on “Body and Soul” and Other Contributions to Music Performance Research, chapter “Body and Soul” and the Mastery of the Jazz Tenor Saxophone.
Chronos.

Ramirez, R., Maestre, E., Pertusa, A., Gómez, E., and Serra, X. (2007).
Performance-based interpreter identification in saxophone audio recordings.
EEE Transactions on Circuits and Systems for Video Technology, 17(3):356–364.

Bibliography VIII

Salamon, J., Peeters, G., and Röbel, A. (2012).
Statistical characterisation of melodic pitch contours and its application for melody extraction.
In *Proceedings of the 13th International Society for Music Information Retrieval Conference (ISMIR)*, pages 187–192, Porto, Portugal.

Serrá, J., Koduri, G. K., Miron, M., and Serra, X. (2011).
Assessing the tuning of sung indian classical music.
In *Proceedings of the 12th International Society for Music Information Retrieval Conference (ISMIR)*, pages 157–162, Miami, USA.

Sundberg, J. (1994).
Acoustic and psychoacoustic aspects of vocal vibrato.
STL-QPSR, 35(2-3):45–68.

Van Herwaarden, S., Grachten, M., and De Haas, W. B. (2014).
Predicting expressive dynamics in piano performances using neural networks.
In *Proceedings of the 15th International Conference on Music Information Retrieval*, Taipei, Taiwan.

Bibliography IX

Widmer, G. and Goebl, W. (2004).
Computational models of expressive music performance: The state of the art.
Journal of New Music Research, 33(3):203–216.

Thank you very much for your attention!

Questions?