SCORE-INFORMED TRACKING AND CONTEXTUAL ANALYSIS OF FUNDAMENTAL FREQUENCY CONTOURS IN TRUMPET AND SAXOPHONE JAZZ SOLOS

Jakob Abeßer
Fraunhofer IDMT
Ilmenau, Germany
Liszt School of Music
Weimar, Germany
jakob.abesser@idmt.fhg.de

Martin Pfleiderer
Liszt School of Music
Weimar, Germany

Klaus Frieler
Liszt School of Music
Weimar, Germany

Wolf-Georg Zaddach
Liszt School of Music
Weimar, Germany

ABSTRACT

In this paper, we propose a novel algorithm for score-informed tracking of the fundamental frequency over the duration of single tones. The tracking algorithm is based on a peak-picking algorithm over spectral magnitudes and ensures time-continuous f_0-curves. From a set of 19 jazz solos from three saxophone and three trumpet players, we collected a set of 6785 f_0-contours in total. We report the results of two exploratory analyses. First, we compared typical contour feature values among different jazz musicians and different instruments. Second, we analyzed correlations between contour features and contextual parameters that describe the metrical position, the in-phrase position, and additional properties of each tone in a solo.

1. INTRODUCTION

The personal style of a musician or singer encompasses various features of her or his performances such as micro-timing, intonation (i.e., pitch accuracy according to a certain tone system), glidings at beginnings and endings or between successive tones and several features of sound, e.g. breathiness or roughness of tones or their overall timbre [1, 2]. However, for the task of Automatic Music Transcription (AMT), tones are commonly understood as acoustic events with a fixed pitch, onset, and offset time [3][1]. This symbolic music representation can be beneficial for a score-level analysis of musical properties such as interval distributions, chords, and scales. At the same time, further artist-specific aspects of a music performance such as pitch glides, intonation, or timbre are completely neglected. Some authors analyze pitch contours as part of automatic melody transcription systems. For instance, Salamon et al. extract different statistical features from pitch contours in polyphonic music, which are used as criteria to assign them to the main melody [4].

When observing jazz improvisation performances of trumpet and saxophone players, the fundamental frequency rarely remains constant over the full duration of a tone. Instead, frequency modulation techniques such as pitch bends, glissandi between tones, vibratos of varying speed and range, and other ornamentations are used to give individual expressiveness to the tones and melodic lines [5]. In African American music genres like jazz, especially thirds, fifths, and sevenths are played in a peculiar way—sometimes a bit too low, sometimes with a gliding movement of the pitch. This phenomenon is often referred to as blue notes or blue note areas by ethnomusicologists [6, 7]. Moreover, jazz musicians often play with vibrato and shape their vibrato in different ways, e.g. faster or slower, or with different amounts of pitch deviations. Depending on jazz style and artist, longer tones are played without vibrato at the beginning and then, often starting on a strong metrical position, are enriched by adding vibrato [8].

In this paper, we primarily focus on the intonation of tones in improvisation of jazz musicians, which could be a pivotal feature of their personal "sound" and playing style. Therefore, we neglect other perceptual aspects related to the instrumental timbre or micro-timing and solely focus on how the fundamental frequency evolves over the duration of a melody tone. In particular, we analyzed audio recordings from six well-known trumpet and saxophone players. We initially extracted tone-wise f_0-contours based on manual transcriptions of instrument solos and then performed two different analyses: Firstly, we investigated whether significant differences can be found for contour feature values among different artists and different instruments. Secondly, we analyzed whether and how contour features such as the deviation of the fundamental frequency from the annotated pitch depend on contextual properties such as the tone's pitch, duration, and metrical position.

This paper is structured as follows. The selection of commercial jazz recordings used for this publications will be described in Section 2.1 In section 2, the proposed algorithm for score-informed f_0-tracking as will be explained in detail. Section 2.7 will give a brief description of the features we extract for each f_0-contour. In Section 3, the differences between artist / instruments and the bias of contextual parameters will be explored. Finally, some conclusions for elaborated transcription strategies and jazz research will be drawn in Section 4.

[1] Throughout this publication, we use the terms *note* for annotated pitch values and *tone* for all sound events that were performed / played on a musical instrument.

2. NOVEL APPROACH

Figure 1 illustrates the proposed method for score-informed estimation of fundamental frequency contours from jazz solos.

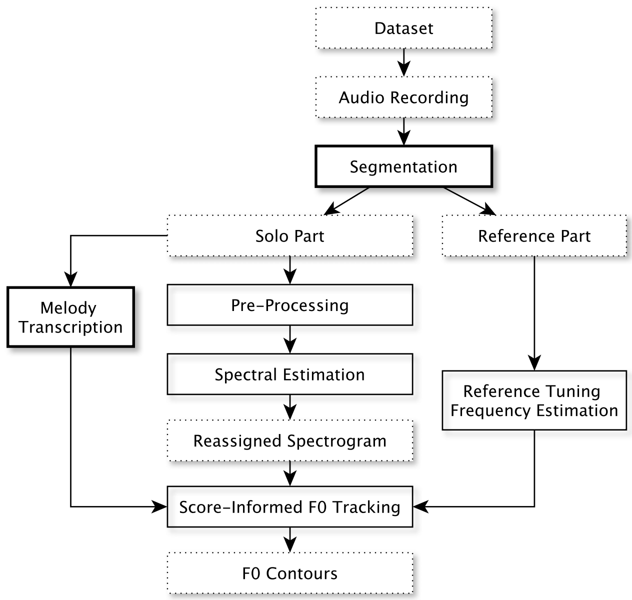

Figure 1: Flow-chart of the proposed algorithm for score-informed f_0-tracking. Steps illustrated with thicker outline are performed manually, all other processing steps are performed automatically.

2.1. Dataset & Melody Transcription

In this publications, we analyze 19 solos played by three saxophone and three trumpet players as listed in Table 1. These solos were taken from the *Weimar Jazz Database* (WJazzD) [9], which currently comprises 174 fully transcribed jazz solos. Based on the original audio recordings, the solos were manually transcribed and cross-checked by musicology and jazz students at the Liszt School of Music. The transcriptions include the MIDI pitch as well as the onset and offset time for each tone played by the soloist. Furthermore, each solo was segmented into melodic phrases, which often coincide with breathing cycles of the saxophone and trumpet players during their improvisation.

2.2. Segmentation

Given a jazz recording, we first manually extract two segments of interest. The *solo part* contains the transcribed instrumental solo and the *reference part* has only the accompanying rhythm section, i.e., piano, double bass, and drums, but no soloist playing. The last two columns of Table 1 show the durations of the solo parts and the reference parts for the used data set.

2.3. Reference Tuning Frequency Estimation

Jazz recordings often show tuning deviations, for instance due to speed variations of tape recorders in the recording process or the particular tuning of the piano used. Amongst others, we aim to analyze the deviation between the note intonation of the soloist and the note intonation of the accompanying rhythm section. Therefore, we first perfom a tuning estimation over the reference part discussed in the previous section to obtain a *reference tuning frequency* f_{ref}.

In particular, we follow the approach implemented in the Chroma Toolbox [10]: Based on a given tuning hypothesis (fundamental frequency of the pitch A4), a triangular filterbank is constructed in such way that its center frequencies are aligned to the semitone fundamental frequencies within the full pitch range of the piano. The STFT magnitude spectrogram is computed, averaged over the duration of the reference part, and filtered using the filterbank to get a measure-of-fit for the current tuning hypothesis. In contrast to the original implementation, we search for the tuning frequency around 440 Hz with a margin of $\pm\frac{1}{2}$ semitone (MIDI pitch range: 69 ± 0.5) and use a very small stepsize of 0.1 cents for the grid search.

2.4. Pre-processing

After the reference tuning frequency f_{ref} is estimated, the next step is to estimate the f_0-contours for each tone played in a given solo. In order to reduce the computation time, we apply a downsampling by factor 2 to a sampling rate of $f_{\text{s}} = 22.05\,\text{kHz}$, since all fudamental frequency values that can be played on the saxophone and the trumpet are below the Nyquist frequency of $f_{\text{s}}/2$. For each tone, the corresponding audio signal is extracted between the tone's onset time and offset time.

2.5. Spectral Estimation

In order to track the fundamental frequency contour over time, we compute a *reassigned magnitude spectrogram* M_{IF} based on the *instantaneous frequency* (IF) as follows. The instantaneous frequency $\hat{f}(k,n)$ for each time-frequency bin in the STFT spectrogram $X(k,n)$ is estimated using the method proposed by Abe in [11]. The approach uses the time derivative of the phase for a frequency correction. We use a zero-padding factor of 16, a blocksize of $b = 2048$, and a hopsize of $h = 64$. The magnitude spectrogram is computed as $M(k,n) = |X(k,n)|$.

We define a logarithmically-spaced frequency axis $f_{\text{log}}(k_{\text{log}})$ with a resolution of 50 bins per semitone. This axis is aligned to the reference tuning frequency f_{ref} and defined for each target tone with a pitch tolerance band of $\pm$ 2 semitones around its annotated MIDI pitch value P as

$$f_{\text{log}}(k_{\text{log}}) = f_{\text{ref}} \cdot 2^{\frac{P-69-2+k_{\text{log}}/50}{12}} \tag{1}$$

The MIDI pitch value of 69 refers to the pitch A4, which corresponds to f_{ref}. The frequency index is denoted as k_{log} with $0 \leq k_{\text{log}} \leq 201$.

The spectral magnitude reassignmend is performed as follows. In each time frame n, each magnitude value $M(k,n)$ (of the STFT magnitude spectrogram) is mapped to the frequency bin $\tilde{k}_{\text{log}}$ of the reassigned spectrogram M_{IF}, whose frequency $f_{\text{log}}(\tilde{k}_{\text{log}})$ is closest to the instantaneous frequency value $\hat{f}(k,n)$. Hence, the original magnitude values of $M(k,n)$ are accumulated in $M_{\text{IF}}(k_{\text{log}},n)$ as follows:

$$M_{\text{IF}}(\tilde{k}_{\text{log}},n) = \sum_k \sum_n \delta(k,n) \cdot M(k,n) \tag{2}$$

Table 1: Selection of saxophone and trumpet solos taken from *Weimar Jazz Database* (WJazzD) that is analyzed in this paper. The columns show the solo number, the artist name, the song title, the solo instrument, the number of notes per solo, the estimated tuning frequency f_{ref} from the reference part (see Section 2.3), the deviation of f_{ref} from the tuning frequency from 440 Hz in cent, as well as the duration of the reference part D_{ref} and the duration of the solo part D_{solo}. The total number of notes per artist is given in brackets after the artist name. The last row shows the mean (μ) and standard deviation (σ) values over all solos. Additional solo metadata can be found at [9].

Solo #	Artist	Title	Instrument	Notes	f_{ref} [Hz]	Δf_{ref} [cent]	D_{ref} [s]	D_{solo} [s]
1	Coleman Hawkins (1195)	Body And Soul	Saxophone	636	445.45	21.3	9.11	167.44
2		My Blue Heaven	Saxophone	213	447.28	28.4	10.58	58.1
3		Stompin' At The Savoy	Saxophone	346	438.71	-5.1	33.86	61.99
4	Michael Brecker (1271)	Midnight Voyage	Saxophone	589	441.94	7.6	33.27	153.86
5		Nothing Personal	Saxophone	682	440.84	3.3	23.82	118.59
6	Sonny Rollins (999)	Blue Seven - 1	Saxophone	354	442.5	9.8	21.51	109.71
7		Blue Seven - 2	Saxophone	138	442.5	9.8	21.51	38.41
8		Tenor Madness	Saxophone	507	438.73	-5	31.06	130.63
9	Clifford Brown (1085)	George's Dilemma	Trumpet	429	440.08	0.3	46.5	100.7
10		Joy Spring	Trumpet	455	441.35	5.3	43.38	94.73
11		Sandu	Trumpet	201	439.59	-1.6	50.68	44.79
12	Freddie Hubbard (1043)	245	Trumpet	481	435.78	-16.7	37.75	120.03
13		Down Under	Trumpet	114	440.28	1.1	25.37	38.56
14		Society Red	Trumpet	448	439.59	-1.6	36.25	149.83
15	Miles Davis (1192)	Blues By Five	Trumpet	371	443.24	12.7	39.49	130.62
16		Oleo - 1	Trumpet	223	442.06	8.1	30.92	56.36
17		Oleo - 2	Trumpet	224	442.06	8.1	30.92	55.26
18		So What	Trumpet	221	452.08	46.9	13.78	112.11
19		Vierd Blues	Trumpet	153	436.81	-12.6	28.62	100.17
μ (σ)				366.95 (213.91)	440.9 (3.24)	3.49 (12.63)	30.98 (9.57)	96.65 (42.38)

with

$$\delta(k,n) = \begin{cases} 1, & \text{if } \tilde{k}_{\text{log}} = \arg\min_{k_{\text{log}}} \left| f_{\text{log}}(k_{\text{log}}) - \hat{f}(k,n) \right| \\ 0, & \text{otherwise.} \end{cases} \tag{3}$$

2.6. Score-Informed f_0-tracking

After computing the reassigned spectrogram $M_{\text{IF}}(k_{\text{log}}, n)$, the f_0-contour of the target tone is tracked over its complete duration. As an example, Figure 2 illustrates a tone taken from the solo "Stompin' At The Savoy" by the saxophonist Coleman Hawkins. The transcibed pitch value is $P = 65$. In the following sections, it will be detailed, how the starting location (indicated as blue circle) is derived and how the f_0-contour (indicated as red circles) is tracked.

2.6.1. Starting Location

Before the f_0-contour can be tracked, a suitable starting location $(k_{\text{log,start}}, n_{\text{start}})$ must be identified. Therefore, we first retrieve the frequency bin positions $k_{\text{log,max}}(n)$ of the frame-wise magnitude maxima as:

$$k_{\text{log,max}}(n) = \arg\max_{k_{\text{log}}} M_{\text{IF}}(k_{\text{log}}, n) \tag{4}$$

Then, we aim to find the frame n_{start}, in which the magnitude peak is closest to the frequency bin $k_{\text{log}} = 100$, which corresponds to the transcribed pitch of the given tone. Therefore, we compute the starting frame for the tracking as

$$n_{\text{start}} = \arg\min_{n} |k_{\text{log,max}}(n) - 100| \tag{5}$$

and set the starting frequency bin to $k_{\text{log,start}} = k_{\text{log,max}}(n_{\text{start}})$. In case multiple frames show a minimum peak distance to the f_0 bin, we select the frame with the highest magnitude in M_{IF}.

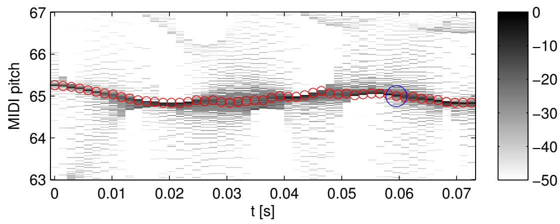

Figure 2: Example f_0-contour of a tone taken from the solo "Stompin' At The Savoy" by the saxophonist Coleman Hawkins with an annotated pitch of $P = 65$. The time axis is normalized such that $t = 0$ refers to the tone onset. The reassigned magnitude spectrogram M_{IF} is shown in dB in the background. The tracked f_0-contour is shown as red circles, the starting location for the forwards-backwards tracking is shown as the bigger blue circle at $t \approx 0.06$ s.

2.6.2. Contour Tracking

After finding the starting location $(k_{\text{log,start}}, n_{\text{start}})$, the f_0-contour is tracked on a frame-wise basis forwards and backwards in time. We assume that the f_0-contours are continuous, hence we only allow a maximum absolute frequency deviation between the fundamental frequency values in adjacent frames of 10 bins, which corresponds to 20 cent for the given frequency axis. In each frame, we choose the f_0 frequency bin based on the maximum peak position in the search range around the previous f_0 estimate. For the backwards tracking, we obtain

$$k_{\text{log},0}(n) = \text{argmax}_{k_{\text{log}}} M_{\text{IF}}(k_{\text{log}}, n) \tag{6}$$
$$\text{for } k_{\text{log},0}(n+1) - 10 \leq k_{\text{log}} \leq k_{\text{log},0}(n+1) + 10$$

The forward tracking is performed in a similar fashion. Hence, the estimated fundamental frequency is $\hat{f}_0(n) = f_{\text{log}}\left(k_{\text{log},0}(n)\right)$.

2.7. Feature Extraction

This section details a set of *contour features*, which are computed to characterize each estimated f_0-contour. We measure the local deviation between the estimated fundamental frequency $\hat{f}_0(n)$ and the annotated fundamental frequency $f_0 = f_{\text{ref}} \cdot 2^{\frac{P-69}{12}}$ in cents as

$$\Delta f_0(n) = 1200 \log_2 \left(\hat{f}_0(n)/f_0 \right). \tag{7}$$

$\Delta f_0(n)$ provides a pitch-independent measure of frequency deviation, which is easy to interpret (100 cents correspond to one semitone). We extract the features

- **AvF0Dev**—median over the frequency deviation $\Delta f_0(n)$, which can indicate a sharp or flat intonation,
- **AvAbsF0Dev**—median over the absolute frequency deviation $|\Delta f_0(n)|$, which measures the total deviation from the reference pitch values,
- **LinF0Slope**—approximated (linear) slope of the f_0-contour over the duration of a tone (based on linear regression over $f_0(n)$),
- **F0Progression**—overall pitch progression in cent from the first to the last 5 % of the tone's total duration [12], and
- three features that can characterize a vibrato by measuring the modulation frequency (**ModFreq**), the total modulation range in cent (**ModRange**), as well as the number of modulation periods (**ModNumPeriod**) [12].

The most prominent modulation frequency detected as from the position of the highest peak of the FFT magnitude spectrogram over $f_0(n)$ in the range between 0.3 and 10 Hz. However, this approach will result in a modulation frequency value for all tones but doesn't necessarily imply a vibrato articulation. As will be discussed in Section 4, future work must adress an initial filtering of tones played with vibrato before the modulation frequency values are further interpreted.

2.8. Contextual Parameters

We want to closer investigate the hypothesis that the intonation of each tone depends on its position in the solo, its position in the current melodic phrase, as well as on its metrical position. Therefore, for each tone in a solo, we extract several *contextual parameters* based on the ground truth transcriptions (see Section 2.1).

While onset time (**Onset**) indicates the position of a tone in the solo in absolute time (seconds), we obtain two features of relative tone position from the melodic phrase annotations:

- **PhraseNum**—number of the corresponding melodic phrase and
- **RelPosInPhrase**—relative position of a tone within that phrase (the relative phrase position is a normalized value with 0 indicating the first tone and 1 indicating the last tone of a melodic phrase).

Besides duration (**Duration**) in seconds, we use two features to indicate the position according to the meter:

- **BeatNum**—corresponding beat number within a bar and
- **SubBeatNum**—corresponding sub-beat number (relates to the tatum, i.e., the metrical subdivision that coincides with most of the tone onsets).

Finally, pitch (**Pitch**) refers to the overall ambitus.

3. STATISTICAL ANALYSIS

In this section, we describe several exploratory analyses that we performed to reveal characteristic correlations and relationships within the data set.

3.1. Feature dependency of artist and instrument

In the first analysis, we investigated the distribution of feature values among solos played by different artists and solos played with different instruments. In particular, we focused on the features **AvF0Dev**, **ModRange**, and **F0Progression**. Figure 3 shows the boxplots over these four features for varying artists and instruments. Several observations can be made:

- Particularly Sonny Rollins and Clifford Brown show a tendency to a sharp intonation with median **AvF0Dev** values of 6.4 and 14.0 cent while Miles Davis tends to a flat intonation (-6.0 cent).
- No strong difference of the **AvF0Dev** feature values can be observed when averaged over all trumpet and saxophone players. This leads to the assumption that the tendency towards a sharp or flat intonation is not instrument-specific but rather artist-specific.
- It can be seen that the saxophone players, especially Coleman Hawkins (median **ModRange** value of 40.8 cent) show a higher modulation range than the trumpet players.
- Concerning the **F0Progression** feature, the results indicate that all of the investigated jazz musicians but Freddy Hubbard show a tendency towards upwards pitch glidings (positive **F0Progression** values). The difference between saxophone and trumpet solos is rather small (5.2 vs. 1.5)

3.2. Correlation between Contour Features and Contextual Parameters

In the second analysis, we investigated the correlations between contour features and contextual parameters. An initial Lilliefors test showed that none of the contour features nor the contextual parameters showed a normal distribution. Therefore, throughout the analyses discussed in this section, we used the Kendall τ rank correlation coefficient. The correlation results between pairs of features and contextual parameters are shown in Table 2 (moderate effect sizes of $|\tau| \geq 0.3$ are emphasized using bold print). The following observations can be made:

- While many highly significant correlations between features and contextual parameters exist (due to the large number of tones), most of them have only a small effect size.
- Neither the tuning deviation **AvF0Dev**, the absolute tuning deviation **AvAbsF0Dev**, the slope of the f_0-contour **LinF0Slope**, nor the **F0Progression** feature seem to depend on investigated contextual parameters.
- With increasing tone duration, the modulation frequency decreases ($\tau = -0.62$) while the modulation range and the number of periods increase ($\tau = 0.23$ and $\tau = 0.32$). Apparently, longer notes are played with slower but more extreme vibrato than shorter notes.
- Also, the modulation range in cent decreases with increasing pitch ($\tau = -0.28$), which is most likely caused by playing difficulties in higher pitch registers. As shown in

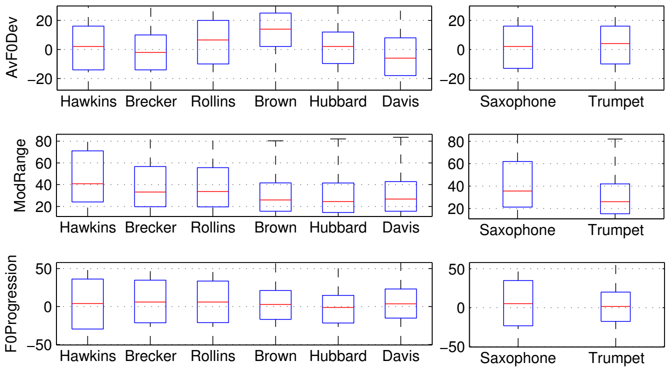

Figure 3: Boxplots over features **AvF0Dev**, **ModRange**, and **F0Progression** over different artists and different instruments.

Table 2: Kendall's τ between features and contextual parameters. Moderate correlation levels ($|\tau| \geq 0.3$) are indicated in bold print. Only significant correlations are shown ($p < .05$). The different significance levels based on the p-value are indicated as *** ($p < .001$), ** ($p < .01$), and * ($p < .05$).

Features	**Contextual Parameters**						
	Metrical Position		**In-phrase Position**		**Basic Tone Parameters**		
	BeatNum	**SubBeatNum**	**PhraseNum**	**RelPosInPhrase**	**Pitch**	**Onset**	**Duration**
AvF0Dev	0.02*	0.02*			-0.07***		-0.04***
AvAbsF0Dev	-0.02*	0.04***		-0.02**	-0.11***		-0.13***
LinF0Slope		-0.03**		-0.03***			-0.02*
ModFreq		0.15***	-0.09***	-0.05***	-0.05***	-0.06***	**-0.62*****
ModRange		-0.05***	0.05***	0.06***	-0.28***	0.05***	0.23***
ModNumPeriod		-0.09***		0.06***			**0.32*****
F0Progression		-0.03***		-0.02**			

the boxplots in Figure 4 and 5, this phenomenon can be observed in a similar fashion for both trumpet and saxophone solos.

4. CONCLUSIONS

In this paper, we propose a novel method for score-informed tracking of fundamental frequency contours. Furthermore, we introduce a set of basic contour features that characterize various aspects such as modulation range, tuning deviation towards the equal-temperament scale, as well as the overall pitch progression. In the second part of our paper, we present several exploratory analyses to investigate, how feature values differ among different artists as well as different instruments and how the contour features and contextual parameters corelate with each other.

This leads to several observations which could be fruitful for future investigations of personal style in jazz improvisation as well as for music research in general. Obviously, different jazz musicians have different tendencies to glide towards or within pitches—with a general trend to glide upwards. Personal vibrato styles are characterized mainly by different modulation ranges as well as minor differences of vibrato frequency. Our method of score-informed tracking of fundamental frequency contours could help to characterize those idiosyncratic vibrato styles.

Preceding to a vibrato analysis, all tones in a solo that are played with vibrato must be identified first. As shown in [12], a supervised classification approach based on contour features such as discussed in Section 2.7 seems as a promising approach. Since vibrato is often put on longer tones (and only occasionally on shorter ones) by jazz musicians, vibrato analyses could be enhanced by filtering the data for longer tones only, which is easily done by the MeloSpySuite software that was developed in the Jazzomat Research Project [13]. Similarly, with MeloSpySuite we could simply filter the data for thirds, fifths, and sevenths according to the underlying chords in order to analyze selectively the pitch contours in those blue note areas and learn about the bias of different artists to play blue notes. In this manner, computer-based methods of transcription and analyses of audio recordings could be extended from the symbolic or structural level (pitch, onset, duration of tones) to the micro-level of musical sound which is, presumably, pivotal for the understanding of performance style in jazz and other music genres.

5. ACKNOWLEDGEMENTS

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 (DFG). The authors would like to thank all jazz and musicology students participating in the transcription and annotation process.

6. REFERENCES

[1] D. Gerhard, "Pitch track target deviation in natural singing," in *Proceedings of the 6th International Conference on Music*

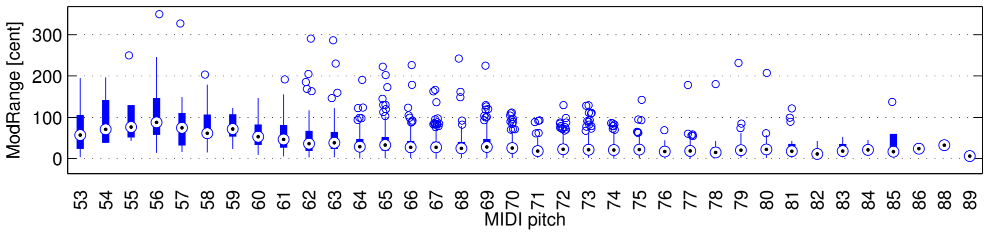

Figure 4: Boxplot over feature **ModRange** in cent over the pitch range for trumpet tones.

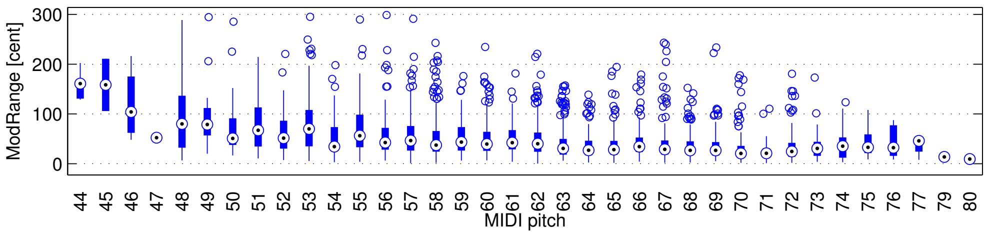

Figure 5: Boxplot over feature **ModRange** in cent over the pitch range for saxophone tones.

Information Retrieval (ISMIR), London, UK, 2005, pp. 514–519.

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

[3] J. Salamon, E. Gómez, D. Ellis, and G. Richard, "Melody extraction from polyphonic music signals: Approaches, applications, and challenges," *Signal Processing Magazine, IEEE*, vol. 31, no. 2, pp. 118–134, March 2014.

[4] J. Salamon, G. Peeters, and A. Röbel, "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)*, Porto, Portugal, 2012, pp. 187–192.

[5] P. F. Berliner, *Thinking in Jazz: The Infinite Art of Improvisation*, Chicago Studies in Ethnomusicology. University of Chicago Press, 1994.

[6] D. Evans, *Big Road Blues: Tradition And Creativity In The Folk Blues*, Da Capo Press, 1987.

[7] J. T. Titon, *Early Downhome Blues. A Musical and Cultural Analysis*, The University of North Carolina Press, 2nd edition, 1995.

[8] M. Pfleiderer, *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, 2009.

[9] "Weimar Jazz Database (WJazzD)," `http://jazzomat.hfm-weimar.de/`, Accessed: 2014-06-09.

[10] M. Müller and S. Ewert, "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)*, 2011, pp. 215–220.

[11] T. Abe, T. Kobayashi, and S. Imai, "Harmonics tracking and pitch extraction based on instantaneous frequency," in *Proceedings of the Int. IEEE Conference on Acoustics, Speech and Signal Processing (ICASSP-95)*, Detroid, USA, May 9 - 12, 1995, pp. 756–759.

[12] J. Abeßer, C. Dittmar, and G. Schuller, "Automatic recognition and parametrization of frequency modulation techniques in bass guitar recordings," in *Proceedings of the 42nd Audio Engineering Society (AES) Conference: Semantic Audio*, Ilmenau, Germany, 2011.

[13] J. Abeßer, K. Frieler, M. Pfleiderer, and W.-G. Zaddach, "Introducing the jazzomat project - jazz solo analysis using music information retrieval methods," in *Proceedings of the 10th International Symposium on Computer Music Multidisciplinary Research (CMMR) Sound, Music and Motion*, Marseille, France, 2013, pp. 187–192.