AUDIO-ALIGNED JAZZ HARMONY DATASET FOR AUTOMATIC CHORD TRANSCRIPTION AND CORPUS-BASED RESEARCH

Vsevolod Eremenko, Emir Demirel, Baris Bozkurt, Xavier Serra
Music Technology Group, Universitat Pompeu Fabra, Barcelona
`firstname.lastname@upf.edu`

ABSTRACT

In this paper we present a new dataset of time-aligned jazz harmony transcriptions. This dataset is a useful resource for content-based analysis, especially for training and evaluating chord transcription algorithms. Most of the available chord transcription datasets only contain annotations for rock and pop, and the characteristics of jazz, such as the extensive use of seventh chords, are not represented. Our dataset consists of annotations of 113 tracks selected from "The Smithsonian Collection of Classic Jazz" and "Jazz: The Smithsonian Anthology," covering a range of performers, subgenres, and historical periods. Annotations were made by a jazz musician and contain information about the meter, structure, and chords for entire audio tracks. We also present evaluation results of this dataset using state-of-the-art chord estimation algorithms that support seventh chords. The dataset is valuable for jazz scholars interested in corpus-based research. To demonstrate this, we extract statistics for symbolic data and chroma features from the audio tracks.

1. INTRODUCTION

Musicians in many genres use an abbreviated notation, known as a lead sheet, to represent chord progressions. Digitized collections of lead sheets are used for computer-aided corpus-based musicological research, e.g., [6, 13, 18, 31]. Lead sheets do not provide information about how specific chords are rendered by musicians [21]. To reflect this rendering, music information retrieval (MIR) and musicology communities have created several datasets of audio recordings annotated with chord progressions. Such collections are used for training and evaluating various MIR algorithms (e.g., Automatic Chord Estimation) and for corpus-based research.

Because providing chord annotations for audio is time-consuming and requires qualified annotators, there are few such datasets available for MIR research. Of the existing datasets, most are of rock and pop music, with very few

available for jazz. A balanced and comprehensive corpus of jazz audio recordings with chord transcription would be a useful resource for developing MIR algorithms aimed to serve jazz scholars. The particularities of jazz also allow us to view the dataset's format, content selection, and chord estimation accuracy evaluation from a different angle.

This paper starts with a review of publicly available datasets that contain information about harmony, such as chord progressions and structural analysis. Based on this review, we justify the necessity of creating a representative, balanced jazz dataset in a new format. We present our dataset, which contains lead sheet style chords, beat onsets, and structure annotations for a selection of jazz audio tracks, along with full-length annotations for each recording. We explain our track selection principle and transcription methodology, and also provide pre-calculated chroma features [27] for the entire dataset. We then discuss how to evaluate the performance of Automatic Chord Transcription on jazz recordings. Moreover, baseline evaluation scores for two state-of-the-art chord estimation algorithms are shown. Dataset is available online [1].

2. RELATED WORKS

2.1 Chord annotated audio datasets

Here we review existing datasets with respect to their format, content selection principle, annotation methodology, and their uses in research. We then discuss some discrepancies in different approaches to chord annotation, as well as the advantages and drawbacks of different formats.

2.1.1 Isophonics family

Isophonics [2] is one of the first time-aligned chord annotation datasets, introduced in [17]. Initially, the dataset consisted of twelve studio albums by The Beatles. Harte justified his selection by stating that it is a small but varied corpus (including various styles, recording techniques and "complex harmonic progressions" in comparison with other popular music artists). These albums are "widely available in most parts of the world" and have had enormous influence on the development of pop music. A number of related theoretical and critical works was also taken into account. Later the corpus was augmented with some

---

[1] `http://doi.org/10.5281/zenodo.1290736`. Documentation: `https://mtg.github.io/JAAH`
[2] Available at `http://isophonics.net/content/reference-annotations`

transcriptions of Carole King, Queen, Michael Jackson, and Zweieck.

The corpus is organized as a directory of ".lab" files[3] . Each line describes a chord segment with a start time, end time (in seconds), and chord label in the "Harte et al." format [16]. The annotator recorded chord start times by tapping keys on a keyboard. The chords were transcribed using published analyses as a starting point, if possible. Notes from the melody line were not included in the chords. The resulting chord progression was verified by synthesizing the audio and playing it alongside the original tracks. The dataset has been used for training and testing chord evaluation algorithms (e.g., for MIREX[4]).

The same format is used for the "Robbie Williams dataset" [5] announced in [12]; for the chord annotations of the RWC and USPop datasets[6] ; and for the datasets by Deng: JayChou29, CNPop20, and JazzGuitar99[7] . Deng presented this dataset in [11], and it is the only one in the family which is related to jazz. However, it uses 99 short, guitar-only pieces recorded for a study book, and thus does not reflect the variety of jazz styles and instrumentations.

2.1.2 Billboard

Authors of the Billboard[8] dataset argued that both musicologists and MIR researchers require a wider range of data [7]. They selected songs randomly from the Billboard "Hot 100" chart in the United States between 1958 and 1991.

Their format is close to the traditional lead sheet: it contains meter, bars, and chord labels for each bar or for particular beats of a bar. Annotations are time-aligned with the audio by the assignment of a timestamp to the start of each "phrase" (usually 4 bars). The "Harte et al." syntax was used for the chord labels (with a few additions to the shorthand system). The authors accompanied the annotations with pre-extracted NNLS Chroma features [27]. At least three persons were involved in making and reconciling a singleton annotation for each track. The corpus is used for training and testing chord evaluation algorithms (e.g., MIREX ACE evaluation) and for musicological research [13].

2.1.3 Rockcorpus and subjectivity dataset

Rockcorpus[9] was announced in [9]. The corpus currently contains 200 songs selected from the "500 Greatest Songs of All Time" list, which was compiled by the writers of *Rolling Stone* magazine, based on polls of 172 "rock stars and leading authorities."

As in the Billboard dataset, the authors specify the structure segmentation and assign chords to bars (and to

beats if necessary), but not directly to time segments. A timestamp is specified for each measure bar.

In contrast to the previous datasets, authors do not use "absolute" chord labels, e.g., `C:maj`. Instead, they specify tonal centers for parts of the composition and chords as Roman numerals. These show the chord quality and the relation of the chord's root to the tonic. This approach facilitates harmony analysis.

Each of the two authors provides annotations for each recording. As opposed to the aforementioned examples, the authors do not aim to produce a single "ground truth" annotation, but keep both versions. Thus it becomes possible to study subjectivity in human annotations of chord changes. The Rockcorpus is used for training and testing chord evaluation algorithms [19], and for musicological research [9].

Concerning the study of subjectivity, we should also mention the Chordify Annotator Subjectivity Dataset[10] , which contains transcriptions of 50 songs from the Billboard dataset by four different annotators [22]. It uses JSON-based JAMS annotation format.

2.2 Jazz-related datasets

Here we review datasets which do not have audio-aligned chord annotations as their primary purpose, but nevertheless can be useful in the context of jazz harmony studies.

2.2.1 Weimar Jazz Database

The main focus of the Weimar Jazz Database (WJazzD[11]) is jazz soloing. Data is disseminated as a SQLite database containing transcription and meta information about 456 instrumental jazz solos from 343 different recordings (more than 132000 beats over 12.5 hours). The database includes: meter, structure segmentation, measures, and beat onsets, along with chord labels in a custom format. However, as stated by Pfleiderer [30], the chords were taken from available lead sheets, "cloned" for all choruses of the solo, and only in some cases transcribed from what was actually played by the rhythm section.

The database's metadata includes the MusicBrainz[12] Identifier, which allows users to link the annotation to a particular audio recording and fetch meta-information about the track from the MusicBrainz server.

Although WJazzD has significant applications for research in the symbolic domain [30], our experience has shown that obtaining audio tracks for analysis and aligning them with the annotations is nontrivial: the MusicBrainz identifiers are sometimes wrong, and are missing for 8% of the tracks. Sometimes WJazzD contains annotations of rare or old releases. In different masterings, the tempo and therefore the beat positions, differs from modern and widely available releases. We matched 14 tracks from WJazzD to tracks in our dataset by the performer's name

---

[3] ASCII plain text files which are used by a variety of popular MIR tools, e.g., Sonic Visualizer [8].

[4] http://www.music-ir.org/mirex/wiki/MIREX_HOME

[5] http://ispg.deib.polimi.it/mir-software.html

[6] https://github.com/tmc323/Chord-Annotations

[7] http://www.tangkk.net/label

[8] http://ddmal.music.mcgill.ca/research/billboard

[9] http://rockcorpus.midside.com/2011_paper.html

[10] https://github.com/chordify/CASD

[11] http://jazzomat.hfm-weimar.de

[12] A community-supported collection of music recording metadata: https://musicbrainz.org

and the date of the recording. In three cases the MusicBrainz Release is missing, and in three cases rare compilations were used as sources. It took some time to discover that three of the tracks ("Embraceable You", "Lester Leaps In", "Work Song") are actually alternative takes, which are officially available only on extended reissues. Beat positions in the other eleven tracks must be shifted and sometimes scaled to match available audio (e.g., for "Walking Shoes"). This may be improved by using an interesting alternative introduced by Balke et al. [3]: a web-based application, "JazzTube," which matches YouTube videos with WJazzD annotations and provides interactive educational visualizations.

2.2.2 Symbolic datasets

The iRb [13] dataset (announced in [6]) contains chord progressions for 1186 jazz standards taken from a popular internet forum for jazz musicians. It lists the composer, lyricist, and year of creation. The data are written in the Humdrum encoding system. The chord data are submitted by anonymous enthusiasts and thus provides a rather modern interpretation of jazz standards. Nevertheless, Broze and Shanahan proved it was useful for corpus-based musicology research: see [6] and [31].

"Charlie Parker's Omnibook data" [14] contains chord progressions, themes, and solo scores for 50 recordings by Charlie Parker. The dataset is stored in MusicXML and introduced in [10].

Granroth-Wilding's "JazzCorpus" [15] contains 76 chord progressions (approximately 3000 chords) annotated with harmonic analyses (i.e., tonal centers and roman numerals for the chords), with the primary goal of training and testing statistical parsing models for determining chord harmonic functions [15].

2.3 Discussion

2.3.1 Some discrepancies in chord annotation approaches in the context of jazz

An article by Harte et al. [16] de facto sets the standard for chord labels in MIR annotations. It describes the basic syntax and a shorthand system. The basic syntax explicitly defines a chord pitch class set. For example, `C:(3, 5, b7)` is interpreted as C, E, G, B♭. The shorthand system contains symbols which resemble chord representations on lead sheets (e.g., `C:7` stands for C dominant seventh). According to [16], `C:7` should be interpreted as `C:(3, 5, b7)`. However, this may not always be the case in jazz. According to theoretical research [25] and educational books, e.g., [23], the 5th degree is omitted quite often in jazz harmony.

Generally speaking, since chord labels emerged in jazz and pop music practice in the 1930s, they provide a higher

level of abstraction than sheet music scores, allowing musicians to improvise their parts [21]. Similarly, a transcriber can use the single chord label `C:7` to mark the whole passage containing the walking bass line and comping piano phrase, without even noticing, "Is the 5th really played?" Thus, for jazz corpus annotation, we suggest accepting the "Harte et al." syntax for the purpose of standardization, but sticking to shorthand system and avoiding a literal interpretation of the labels.

There are two different approaches to chord annotation:

- "Lead sheet style." Contains a lead sheet [21], which has obvious meaning to musicians practicing the corresponding style (e.g., jazz or rock). It is aligned to audio with timestamps for beats or measure bars. Chords are considered in a rhythmical framework. This style is convenient because the annotation process can be split into two parts: lead sheet transcription done by a qualified musician, and beats annotation done by a less skilled person or sometimes even automatically performed.

- "Isophonics style." Chord labels are bound to absolute time segments.

We must note that musicians use chord labels for instructing and describing performance mostly within the lead sheet framework. While the lead sheet format and the chord-beats relationship is obvious, detecting and interpreting "chord onset" times in jazz is an unclear task. The widely used comping approach to accompaniment [23] assumes playing phrases instead of long isolated chords, and a given phrase does not necessarily start with a chord tone. Furthermore, individual players in the rhythm section (e.g., bassist and guitarist) may choose different strategies: they may anticipate a new chord, play it on the downbeat, or delay. Thus, before annotating "chord onset" times, we should make sure that it makes musical and perceptual sense. All known corpus-based research is based on "lead sheet style" annotated datasets. Taking all these considerations into account, we prefer to use the lead sheet approach to chord annotations.

2.3.2 Criteria for format and dataset for chord annotated audio

Based on the given review and our own hands-on experience with chord estimation algorithm evaluation, we present our guidelines and propositions for building an audio-aligned chord dataset.

1. Clearly define dataset boundaries (e.g., a certain music style or time period). The selection of audio tracks should be representative and balanced within these boundaries.

2. Since sharing audio is restricted by copyright laws, use recent releases and existing compilations to facilitate access to dataset audio.

3. Use the time-aligned lead sheet approach with "shorthand" chord labels from [16], but avoid their literal interpretation.

---

[13] https://musiccog.ohio-state.edu/home/index.php/iRb_Jazz_Corpus
[14] https://members.loria.fr/KDeguernel/omnibook/
[15] http://jazzparser.granroth-wilding.co.uk/JazzCorpus.html

4. Annotate entire tracks, but not excerpts. This makes it possible to explore structure and self-similarity.

5. Provide the MusicBrainz identifier to exploit meta-information from this service. If feasible, add meta-information to MusicBrainz instead of storing it privately within the dataset.

6. Annotate in a format that is not only machine readable, but convenient for further manual editing and verification. Relying on plain text files and specific directory structure for storing heterogeneous annotation is not practical for users. JSON-based JAMS format introduced by Humphrey et al. [20] solves this issue, but currently does not support lead sheet chord annotation. It is verbose in order to be comfortably used by the human annotators and supervisors.

7. Include pre-extracted chroma features. This makes it possible to conduct some MIR experiments without accessing the audio. It would be interesting to incorporate chroma features into corpus-based research to demonstrate how a particular chord class is rendered in a particular recording.

3. PROPOSED DATASET

3.1 Data format and annotation attributes

Taking into consideration the discussion from the previous section, we decided to use the JSON format. An excerpt from an annotation is shown in Figure 1. We provide the track title, artist name, and MusicBrainz ID. The start time, duration of the annotated region, and tuning frequency estimated automatically by Essentia [5] are shown. The beat onsets array and chord annotations are nested into the "parts" attribute, which in turn could recursively contain "parts." This hierarchy represents the structure of the musical piece. Each part has a "name" attribute which describes the purpose of the part, such as "intro," "head," "coda," "outro," "interlude," etc. The inner form of the chorus (e.g., AABA, ABAC, blues) and predominant instrumentation (e.g., ensemble, trumpet solo, vocals female, etc.) are annotated explicitly. This structural annotation is beneficial for extracting statistical information regarding the type of chorus present in the dataset, as well as other musically important properties. We made chord annotations in the lead sheet style: each annotation string represents a sequence of measure bars, delimited with pipes: "|". A sequence starts and ends with a pipe as well. Chords must be specified for each beat in a bar (e.g., four chords for 4/4 meter). A simplification of this is possible: if a chord occupies the whole bar, it could be typed only once; and if chords occupy an equal number of beats in a bar (e.g., two beats in 4/4 metre), each chord could be specified only once, e.g., `|F G|` instead of `|F F G G|`.

For chord labeling, we use the Harte et al. [16] syntax for standardization reasons, but mainly use the shorthand system and do not assume the literal interpretation of labels

Figure 1. An annotation example.

as pitch class sets. More details on chord label interpretation will follow in 4.1.

3.2 Content selection

The community of listeners, musicians, teachers, critics and academic scholars defines the jazz genre, so we decided to annotate a selection chosen by experts.

After considering several lists of seminal recordings compiled by authorities in jazz history and in musical education [14, 24], we decided to start with "The Smithsonian Collection of Classic Jazz" [1] and "Jazz: The Smithsonian Anthology" [2].

The "Collection" was compiled by Martin Williams and first issued in 1973. Since then, it has been widely used for jazz history education and numerous musicological research studies draw examples from it [26]. The "Anthology" contains more modern material compared to the "Collection." To obtain unbiased and representative selection, its curators used a multi-step polling and negotiation process involving more than 50 "jazz experts, educators, authors, broadcasters, and performers." Last but not least, audio recordings from these lists can be conveniently obtained: each of the collections are issued in a CD box.

We decided to limit the first version of our dataset to jazz styles developed before free jazz and modal jazz, because lead sheets with chord labels cannot be used effectively to instruct or describe performances in these latter styles. We also decided to postpone annotating compositions which include elements of modern harmonic structures (i.e., modal or quartal harmony).

3.3 Transcription methodology

We use the following semi-automatic routine for beat detection: the DBNBeatTracker algorithm from the madmom package is run [4]; estimated beats are visualized and sonified with Sonic Visualizer; if needed, DBNBeatTracker is re-run with a different set of parameters; and finally beat annotations are manually corrected, which is usually necessary for ritardando or rubato sections in a performance.

After that, chords are transcribed. The annotator aims to notate which chords are played by the rhythm section.
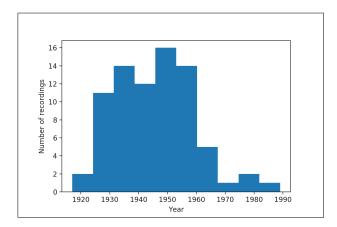
Figure 2. Distribution of recordings from the dataset by year.

If the chords played by the rhythm section are not clearly audible during a solo, chords played in the "head" are replicated. Useful guidelines on chord transcription in jazz are given in the introduction of Henry Martin's book [26]. The annotators used existing resources as a starting point, such as published transcriptions of a particular performance or Real book chord progressions, but the final decisions for each recording were made by the annotator. We developed an automation tool for checking the annotation syntax and chord sonification: chord sounds are generated with Shepard tones and mixed with the original audio track, taking its volume into account. If annotation errors are found during syntax check or while listening to the sonification playback, they are corrected and the loop is repeated.

4. DATASET SUMMARY AND IMPLICATIONS FOR CORPUS BASED RESEARCH

To date, 113 tracks are annotated with an overall duration of almost 7 hours, or 68570 beats. Annotated recordings were made from music created between 1917 and 1989, with the greatest number coming from the formative years of jazz: the 1920s-1960s (see Figure 2). Styles vary from blues and ragtime to New Orleans, swing, be-bop and hard bop with a few examples of gypsy jazz, bossa nova, Afro-Cuban jazz, cool, and West Coast. Instrumentation varies from solo piano to jazz combos and to big bands.

4.1 Classifying chords in the jazz way

In total, 59 distinct chord classes appear in the annotations (89, if we count chord inversions). To manage such a diversity of chords, we suggest classifying chords as it done in jazz pedagogical and theoretical literature. According to the article by Strunk [32], chord inversions are not important in the analysis of jazz performance, perhaps because of the improvisational nature of bass lines. Inversions are used in lead sheets mainly to emphasize the composed bass line (e.g., pedal point or chromaticism). Therefore, we ignore inversions in our analysis.

According to numerous instructional books, and to theoretical work done by Martin [25], there are only five main
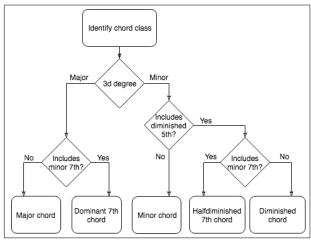

Figure 3. Flow chart: how to identify chord class by degree set.

Chord class	Beats Number	Beats %	Duration (seconds)	Duration %
dom7	29786	43.44	10557	42.23
maj	18591	27.11	6606	26.42
min	13172	19.21	4681	18.72
dim	1677	2.45	583	2.33
hdim7	1280	1.87	511	2.04
no chord	3986	5.81	2032	8.13
unclassified	78	0.11	30	0.12

Table 1. Chord classes distribution.

chord classes in jazz: major (maj), minor (min), dominant seventh (dom7), half-diminished seventh (hdim7), and diminished (dim). Seventh chords are more prevalent than triads, although sixth chords are popular in some styles (e.g., gypsy jazz). Third, fifth and seventh degrees are used to classify chords in a bit of an asymmetric manner: the unaltered fifth could be omitted in the major, minor and dominant seventh (see chapter on three note voicing in [23]); the diminished fifth is required in half-diminished and in diminished chords; and $\flat\flat7$ is characteristic for diminished chords. We summarize this classification approach in the flow chart in Figure 3.

The frequencies of different chord classes in our corpus are presented in Table 1. The dominant seventh is the most popular chord, followed by major, minor, diminished and half-diminished. Chord popularity ranks differ from those calculated in [6] for the iRb corpus: dom7, min, maj, hdim, and dim. This could be explained by the fact that our dataset is shifted toward the earlier years of jazz development, when major keys were more pervasive.

4.2 Exploring symbolic data

Exploring the distribution of chord transition bigrams and n-grams allows us to find regularities in chord progressions. The term bigram for two-chord transitions was de-
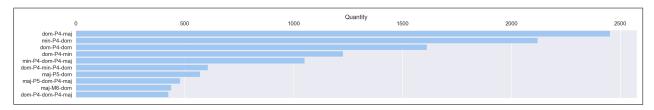
Figure 4. Top ten chord transition n-grams. Each n-gram is expressed as sequence of chord classes (dom, maj, min, hdim7, dim) alternated with intervals (e.g., P4 - perfect fourth, M6 - major sixth), separating adjacent chord roots.

fined in [6]. Similarly, we define an n-gram as a sequence of n chord transitions. The ten most frequent n-grams from our dataset are presented in Figure 4. The picture presented by the plot is what would be expected for a jazz corpus: we see the prevalence of the root movement by the cycle of fifths. The famous IIm-V7-I three-chord pattern (e.g., [25]) is ranked number 5, which is even higher than most of the shorter two-chord patterns.

5. CHORD TRANSCRIPTION ALGORITHMS BASELINE EVALUATION

Now we turn to Automatic Chord Estimation (ACE) evaluation for jazz. We adopt the MIREX [16] approach to evaluating ACE algorithms. The approach supports multiple ways to match ground truth chord labels with predicted labels, by employing the different chord vocabularies introduced by Pauwels [29]. The distinctions between the five chord classes defined in 4.1 are crucial for analyzing jazz performance. More detailed transcriptions (e.g., a distinction between maj6 and maj7, detecting extensions of dom7, etc.) are also important but secondary to classification into the basic five classes. To formally implement this concept of chord classification, we develop a new vocabulary, called "Jazz5," which converts chords into the five classes according to the flowchart in Figure 3.

For comparison, we also choose two existing MIREX vocabularies: "Sevenths" and "Tetrads," because they ignore inversions and can distinguish between major, minor and dom7 classes (which together occupy about 90% of our dataset). However, these vocabularies penalize differences within a single basic class (e.g., between a major triad and a major seventh chord). Moreover, the "Sevenths" vocabulary is too basic; it excludes a significant number of chords, such as diminished chords or sixths, from evaluation.

We choose Chordino [17], which has been a baseline algorithm for the MIREX challenge over several years, and CREMA [18], which was recently introduced in [28]. To date, CREMA is one of the few open-source, state-of-the-art algorithms which supports seventh chords.

Results are provided in the Table 2. "Coverage" signifies the percentage of the dataset which can be evaluated using the given vocabulary. "Accuracy" stands for the per-

Vocabulary	Coverage %	Chordino Accuracy %	CREMA Accuracy %
"Jazz5"	99.88	32.68	40.26
MirexSevenths	86.12	24.57	37.54
Tetrads	99.90	23.10	34.30

Table 2. Comparison of coverage and accuracy evaluation for different chord dictionaries and algorithms.

centage of the covered dataset for which chords were properly predicted, according to the given vocabulary.

We see that the accuracy for the jazz dataset is almost half of the accuracy achieved by the most advanced algorithms on datasets currently involved in the MIREX challenge [19] (which is roughly 70-80%). Nevertheless, the more recent algorithm (CREMA) performs significantly better than the old one (Chordino) which shows that our dataset passes a sanity check: it does not contradict technological progress in Automatic Chord Estimation. We see from this analysis that the "Sevenths" chords vocabulary is not appropriate for a jazz corpus because it ignores almost 14% of the data. We also note that the "Tetrads" vocabulary is too punitive: it penalizes up to 9% of predictions. However, this could potentially be tolerable in the context of jazz harmony analysis. We provide code for this evaluation in the project repository.

6. CONCLUSIONS AND FURTHER WORK

We have introduced a dataset of time-aligned jazz harmony transcriptions, which is useful for MIR research and corpus-based musicology. We have demonstrated how the particularities of the jazz genre affect our approach to data selection, annotation, and evaluation of chord estimation algorithms.

Further work includes growing the dataset by expanding the set of annotated tracks and adding new features. Functional harmony annotation (or local tonal centers) is of particular interest, because we could then implement chord detection accuracy evaluation based on jazz chord substitution rules.

---

[16] http://www.music-ir.org/mirex/wiki/2017:Audio_Chord_Estimation
[17] http://www.isophonics.net/nnls-chroma
[18] https://github.com/bmcfee/crema

[19] http://www.music-ir.org/mirex/wiki/2017:Audio_Chord_Estimation_Results

8. REFERENCES

[1] The Smithsonian Collection of Classic Jazz. Smithsonian Folkways Recording, 1997.

[2] Jazz: The Smithsonian Anthology. Smithsonian Folkways Recording, 2010.

[3] Stefan Balke, Christian Dittmar, Jakob Abeßer, Klaus Frieler, Martin Pfleiderer, and Meinard Müller. Bridging the Gap: Enriching YouTube Videos with Jazz Music Annotations. *Frontiers in Digital Humanities*, 5, 2018.

[4] Sebastian Böck, Filip Korzeniowski, Jan Schlüter, Florian Krebs, and Gerhard Widmer. madmom: a new Python Audio and Music Signal Processing Library. In *Proc. of the 24th ACM International Conference on Multimedia*, pages 1174–1178, 2016.

[5] Dmitry Bogdanov, Nicolas Wack, Emilia Gómez, Sankalp Gulati, Perfecto Herrera, O. Mayor, Gerard Roma, Justin Salamon, J. R. Zapata, and Xavier Serra. Essentia: an audio analysis library for music information retrieval. In *Proc. of the of the International Society for Music Information Retrieval Conference (ISMIR)*, pages 493–498, 2013.

[6] Yuri Broze and Daniel Shanahan. Diachronic Changes in Jazz Harmony: A Cognitive Perspective. *Music Perception: An Interdisciplinary Journal*, 3(1):32–45, 2013.

[7] John Ashley Burgoyne, Jonathan Wild, and Ichiro Fujinaga. An Expert Ground-Truth Set for Audio Chord Recognition and Music Analysis. In *Proc. of the of the International Society for Music Information Retrieval Conference (ISMIR)*, number ISMIR, pages 633–638, 2011.

[8] Chris Cannam, Christian Landone, Mark Sandler, and Juan Pablo Bello. The Sonic Visualiser: A Visualisation Platform for Semantic Descriptors from Musical Signals. In *Proc. of the of the International Society for Music Information Retrieval Conference (ISMIR)*, pages 324–327, 2006.

[9] Trevor de Clercq and David Temperley. A corpus analysis of rock harmony. *Popular Music*, 30(01):47–70, jan 2011.

[10] Ken Déguernel, Emmanuel Vincent, and Gérard Assayag. Using Multidimensional Sequences For Improvisation In The OMax Paradigm. In *13th Sound and Music Computing Conference*, 2016.

[11] Junqi Deng and Yu-kwong Kwok. A hybrid gaussian-hmm-deep-learning approach for automatic chord estimation with very large vocabulary. In *Proc. 17th International Society for Music Information Retrieval Conference*, pages 812–818, 2016.

[12] Bruno Di Giorgi, Massimiliano Zanoni, Augusto Sarti, and Stefano Tubaro. Automatic chord recognition based on the probabilistic modeling of diatonic modal harmony. In *Proc. of the 8th International Workshop on Multidimensional Systems (nDS)*, pages 1–6, 2013.

[13] Hubert Léveillé Gauvin. " The Times They Were A-Changin' " : A Database-Driven Approach to the Evolution of Harmonic Syntax in Popular Music from the 1960s. *Empirical Musicology Review*, 10(3):215–238, apr 2015.

[14] Ted Gioia. *The Jazz Standards: A Guide to the Repertoire*. Oxford University Press, 2012.

[15] Mark Granroth-Wilding. *Harmonic analysis of music using combinatory categorial grammar*. PhD thesis, University of Edinburgh, 2013.

[16] C Harte, M Sandler, S Abdallah, and E Gómez. Symbolic representation of musical chords: A proposed syntax for text annotations. In *Proc. of the of the International Society for Music Information Retrieval Conference (ISMIR)*, volume 56, pages 66–71, 2005.

[17] Christopher Harte. *Towards automatic extraction of harmony information from music signals*. PhD thesis, Queen Mary, University of London, 2010.

[18] Thomas Hedges, Pierre Roy, and François Pachet. Predicting the Composer and Style of Jazz Chord Progressions. *Journal of New Music Research*, 43(3):276–290, jul 2014.

[19] Eric J Humphrey and Juan P Bello. Four timely insights on automatic chord estimation. In *Proc. of the of the International Society for Music Information Retrieval Conference (ISMIR)*, 2015.

[20] Eric J Humphrey, Justin Salamon, Oriol Nieto, Jon Forsyth, Rachel M Bittner, and Juan P Bello. Jams: a Json Annotated Music Specification for Reproducible Mir Research. In *Proc. of the International Society for Music Information Retrieval (ISMIR)*, 2014.

[21] Barry Dean Kernfeld. *The story of fake books : bootlegging songs to musicians*. Scarecrow Press, 2006.

[22] Hendrik Vincent Koops, Bas de Haas, John Ashley Burgoyne, Jeroen Bransen, and Anja Volk. Harmonic subjectivity in popular music. Technical Report UU-CS-2017-018, Department of Information and Computing Sciences, Utrecht University, 2017.

[23] Mark Levine. *The Jazz Piano Book*. Sher Music, 1989.

[24] Mark Levine. *The Jazz Theory Book*. Sher Music, 2011.

[25] Henry Martin. Jazz Harmony: A Syntactic Background. *Annual Review of Jazz Studies*, 8:9–30, 1988.

[26] Henry Martin. *Charlie Parker and Thematic Improvisation*. Institute of Jazz Studies, Rutgers–The State University of New Jersey, 1996.

[27] Matthias Mauch and Simon Dixon. Approximate note transcription for the improved identification of difficult chords. In *Proc. of the of the International Society for Music Information Retrieval Conference (ISMIR)*, number 1, pages 135–140, 2010.

[28] Brian Mcfee and Juan Pablo Bello. Structured Training for Large-Vocabulary Chord Recognition. In *Proc. of the International Conference on Music Information Retrieval (ISMIR)*, pages 188–194, 2017.

[29] Johan Pauwels and Geoffroy Peeters. Evaluating automatically estimated chord sequences. In *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, pages 749–753. IEEE, 2013.

[30] Martin Pfleiderer, Klaus Frieler, and Jakob Abeßer. *Inside the Jazzomat - New perspectives for jazz research*. Schott Campus, Mainz, Germany, 2017.

[31] Keith Salley and Daniel T. Shanahan. Phrase Rhythm in Standard Jazz Repertoire: A Taxonomy and Corpus Study. *Journal of Jazz Studies*, 11(1):1, nov 2016.

[32] Steven Strunk. Harmony (i). The New Grove Dictionary of Jazz, pp. 485–496, 1994.