

機能と声解析データの作成とその統計解析

金子仁美^{†1} 川上大輔^{†2} 嵯峨山 茂樹^{†2}

我々は、楽曲の和声解析の記述仕様 (“KS notation”) を策定し、機能と声解析を行ってデータを作成し、その統計解析を行った。和声推定は自動採譜や楽曲検索など多数の目的に有用で、その和声進行の確率モデルの作成と統計学習のために有用である。また、音楽学的な見地からは、和声学の規則や傾向などが計量的に検証でき、時代や作曲家や楽曲スタイルを和声学的に解明する基礎となる。機能と声記述のために、和音、転回、借用和音、省略、変位、転調、付加音などの記述を可能とし、さらに楽譜なしで演奏が可能ないように音価も表現した。また、人間とコンピュータ双方の可読性の両立させコンパクトに表現できるようにした。データ作成には、RWC 音楽データベース所収のクラシック曲 50 曲について、人手により機能と声解析してデータを作成した。そのデータを統計解析し、音楽的な知見から説明を試み、機能と声モデルが従来のモデルより工学的和声モデルとして優位であることを示す。

Development of functional harmony labels and statistical analysis of chord progressions

HITOMI KANEKO,^{†1} DAISUKE KAWAKAMI^{†2}
and SHIGEKI SAGAMAYA^{†2}

We designed a new notation (called “KS notation”) for harmony analysis, built a functional harmony analysis dataset and made statistical analyses on the data. Harmony (chord sequence) estimation is useful in many purposes including automatic music transcription and music information retrieval, while, from musicological viewpoint, harmony theory and rules are verified quantitatively using the data across periods, composers and styles may be investigated. For description of functional harmony analysis, the notation include chord, inversion, borrowed chord, omission, alteration, key modulation, additional notes, etc. and enables playing chords from the notation without the score by representing the note value. Readability was emphasized both for human and computer. The KS-notation dataset was built from 50 titles included in the RWC classical music database. New findings are discussed based on statistical analysis of the data and functional harmony model is shown to be advantageous over the conventional chord sequence model from the engineering point of view.

1. はじめに

調が崩壊する 20 世紀初頭までの間、和声は西洋音楽の重要な要素であった。調は、芸術音楽変革の歴史の中では 100 年近く前に衰退したが、しかし現代の実生活で人々が耳にし、口ずさみ、愛聴する音楽の大多数は調性による音楽であろう。我々が最も身近に接している、その調性音楽の楽曲を解釈するには、和声を解析することが必須であり、それは楽曲構造やスタイル、作曲の技法の理解、作曲家の特質を判別する上でも意義があり、音楽音響信号や MIDI 信号や楽譜情報などから和声構造を推定することができれば、自動伴奏生成、自動採譜への前処理、同一曲の検索、曲のジャンルやムード分類、自動編曲など多くの用途が考えられる。

我々の研究グループは 1999 年および 2000 年に HMM (Hidden Markov Model) による確率的な和声モデル^{(1),(2)} を提案した。以後、音楽音響信号入力からの和声推定においても確率的な和声モデルがこの分野の主要なアプローチになっている^{(3)-(7),(10),(11)}。

これには、図 1 左のようなただ和音の系列をエルゴディック HMM で表現したモデルや、右のような和音系列を語彙として捉えそれらのネットワークが調を構成し、調の間の遷移 (転調) も含めたモデルまで、様々なモデルが含まれた。本論文の一つの目的は、調を考慮したモデルが音楽理論に合致し、かつ工学的な音楽モデルとしての妥当性を示すことである^{(8),(9)}。

確率的な和声モデルにおいては、和声の出現や進行の統計的学習が必要である。そのためには音響信号あるいは楽譜に和声ラベルが付された学習データが必要である。C. Harte らは、The Beatles による音源の演奏時間を区間として和音のラベルを付与した⁽¹²⁾。しかし、これは現象としての和音のラベルづけであり、同じ構造の和声でも調が異なれば異なった表記がなされており、調と和声の音楽理論に基づいた構造の把握にはなっていなかった。

筆者らは、和声に関する研究推進のため、人手による和声ラベル作業の容易さと、コンピュータ可読性の両立を主眼にして、和声記述仕様を策定し、それに基づいて機能と声ラベルを付与したデータを作成し、その統計解析を行った。本論文では、主に機能と声ラベルとその記述仕様について述べる。

†1 桐朋学園大学音楽学部
Toho Gakuen School of Music

†2 東京大学大学院情報理工学系研究科
Graduate School of Information Science and Technology, The University of Tokyo

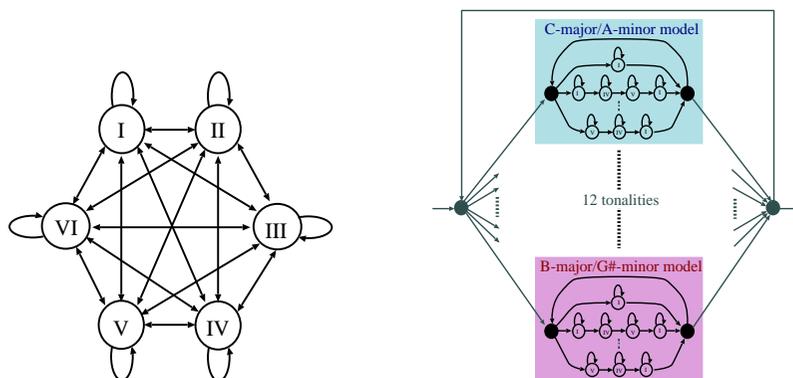


図1 エルゴディック HMM に基づく和音系列モデル (左) と和音系列語彙と転調を含む HMM 和声モデル (右)²⁾

2. 機能と声解析と表記

2.1 機能と声解析とは

音楽にはさまざまなジャンルがあるが、現在日常的に親しまれている音楽の多くは調性音楽、または調性要素を多く含む音楽である。調性音楽の三大要素である旋律、リズム、和声の中で、和声は調と最も密接に係わる要素である。なぜなら和声とは、和音の形態、各調における和音間の進行を解釈することであり、各調でそれぞれに役割を持っているからである。例えば、同じ音名による和音が、調が変わることで役割と進行を変えてしまうということが起きる。

図2にその例を挙げる。左の4つの和音は、同一の和音であるが、想定している調によりその機能が異なり、和音の根音を各調の主音から数えた音程により和音度として論じる。これらは和音としては同一であるが、和声としては異なる、と考えることができる。一方、同じ和音度であっても、調により実際の和音が異なる。すなわち、和声的には同一であっても、和音としては別の実現がなされる。

実は、よく知られている「調」の概念自体、機能と声理論に基づいている。「ある楽曲が xx 調である」ということは、その曲の和音が xx 調に基づく機能と声理論に適合しているということを表している。これが単に理論でなく体感できることから、機能と声理論を知らない人にも調は感じられ判断できる。また、そうであることを理論付けたものを広い意味の機能と声理論と言うことができる。

このような和声の意味を内包し、同時に、その場ごとの和音形態の表現も可能にするのが

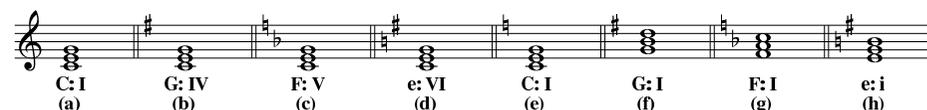


図2 調と和音の関係 — (a)–(d): 和音が調が変わることによって変わる。(e)–(h): 調が変わっても和音記号が変わらない

「機能による和声表記」である。クラシック、ポピュラー、ジャズなど各ジャンルで複数の和声表記が存在する中、和声を調との関係によって認識、解析する機能による和声表記は、工学研究、音楽学研究において有益と考えられる（更には非調性音楽に対しても、和音の機能が有効か無効かを認識することで、楽曲の特徴を解析できると期待できる。）

2.2 機能と声解析と表記

機能による和声解析とは、端的に言えば、調判別をし、その調の主音と和音の根音の距離（音程）をローマ数字で表現することである。このローマ数字が和音の度数（和音度）と言われ、和音の個々の形態を表現するだけでなく、その調における役割（機能）を明確化する。表記は各調ごとの、主音との音程関係によるため、相対的であることが大きな特徴で、転調した場合でも同じ記号により和音を記述できる、という利点がある。更に、2つの和音の連なりのパターンを単位とした解釈を実現し、統計解析にも有用と考えられる。

各機能にはそれぞれの特徴がある。T(tonic) は、その調の主音を根音とする和音で、「緩和」或は「解決」をもたらす。D(dominant) は、その調の属音を根音とする和音で、「緊張」をもたらす。「緩和」「解決」(T) に向かう。S(subdominant) は、ドミナントを経過してトニックに向かうケースと直接トニックに向かうケースがあり、その役割、用法は T, D と比較すると複数の可能性を持つ。和音の連なりはカデンツを形成する。カデンツには、[1]T D T, [2]T S D T, [3]T S T という3種の形態があり、調性音楽のほとんどの和音連鎖は、この形態に集約される。カデンツは調性音楽でのみ意義を持つ和音連鎖パターンである。調性音楽になり得るにはカデンツによる調の確定が必要不可欠である。

2.3 伝統的記法

和声表記法は、音楽史上、現在に至るまで複数存在し、国際的に統一されていない。また、研究目的の観点から、主要な従来の和声の表記法にはいくつかの問題がある。

図3に代表的な記号法の例を示す。例えば、Paul Hindemith, Arnold Schoenberg らによる機能と声の表記には和音種別（短三和音と長三和音など）の区別がなく、和音変位の詳細も区別されない。日本で広く使われている機能と声表記¹⁴⁾でも、和音の種別が和音度記号に表現されない。しばしば「楽譜を見れば明らかだから」と説明されるが、それでは計算機処理に馴染まない。また、これらは記号が複雑であり、テキストデータとしてのラベルデータ作成、及び計算機による処理には難がある。

	traditional notation	ii_3^4		traditional notation	$IV_{5/3}^{6+}$
	Shimaoka's notation	II_7^2		Shimaoka's notation	$\circ V_9^2$
	KS notation	$ii7''$		KS notation	$V:V-9''$

図3 各種の和声記法の比較 - 伝統記法, 日本式(芸大和声), KS notation / 左: 八長調 2 度音第二転回形 / 右: 八長調の属調の属九和音の第二転回の根音省略と五度音と九度音の半音下方変位形 (いわゆるドイツ六度和音)

図4 に一般的に用いられる様々な和声表記と, 本論文が提案する KSN との比較を例示する. この例は, 前半の 4 小節と後半の 4 小節は調が異なるが和声進行は基本的に同じであり, 解析には機能と声的表記が不可欠である好例である. 典型的なポピュラー和声の記法ではこの繰り返し構造が表現しにくい. ポピュラー和声は, F, Am, Gdim7, 等のように和声の構成音に関する情報が含まれず演奏者のための和声表記といえる. 調が不明確であり, 転調の位置も不明確である.

2.4 新しい和声記述仕様の必要性

前節で述べた理由から, 和声記述には以下の情報が必要とされる. これらの情報を含む表記は, 和音を機能的に捉え, かつ原曲の各和音情報の詳細も含む.

- 調性 (転調)
- 和音度
- 転回位置
- 和音の種類 (長三和音, 減三和音など)
- 変化音, 付加音
- 省略音

これらの情報は, 調判別を前提としていることから, ローマ数字の和音度表記することが必要とされる. 逆に, 非和声音, 各音のオクターブ位置, 強弱や奏法やパートの区別はしない. なぜなら, これらまで与えれば楽譜と同等になり, 和声的視点とは異なるからである.

3. 機能による和声ラベルデータの構築

3.1 KS notation の設計

従来の和声表記に準じたラベルデータの問題点を解消するため, 次の点に留意して和声ラベルデータの表記法 (KS notation: KSN) を定めた.

c)	5 3	6 4	5 4	6 4	6 4	5 3	6# 4	5 3	6# 4	6# 4	7 #										
d)	IV(A):I	V ₃ ⁴	i	V ₃ ⁴	V(B):ii ^ø ₃ ⁴	V ₇	I	V ₃ ⁴	i	V ₃ ⁴	VI(C#):ii ^ø ₃ ⁴	V ₇									
e)	IV(A):I	V ₇ ²	i	V ₇ ²	V(B):ii ^ø ₇ ²	V ₇	I	V ₇ ²	i	V ₇ ²	VI(°ii ^ø ₇ ²	V ₇									
f)	A	E _{7/B} (E ₇ ^{on B})	Am	E _{7/B}	C#7-5/G	F#7	B	F#7/C#	Bm	F#7/C#	D#m7-5/A	G#7									
g)	A	E _{7/B}	A-	E _{7/B}	C# ^ø /G (C# ^ø _{m7/G} or E ₆ /G)	F#7	B	F# ₇ /C#	B-	F# ₇ /C#	D# ^ø /A (D# ^ø _{m7/A} or F# ₆ /G)	G#7									
h)	{ IV: I	V7''		i	V7''	}	{ V: ii-7''		ii-7''	V7		I	V7''		i	V7''	}	{ VI: ii-7''		ii-7''	V7

- (a) 楽譜 (F. Chopin: Etude Op.10, No.3, 30-37 小節)
 (b) 和声譜
 (c) 数字つき低音 (figured bass)
 (d) クラシックローマ数字表記 (classical roman)
 (e) 日本式表記 (島岡ら¹⁴⁾ による, いわゆる「芸大和声」)
 (f) 典型的なポピュラー和声表記
 (g) 典型的なジャズ和声表記
 (h) KS 表記 (提案表記法)

図4 各種の和声記法の比較 (Chopin Op.10, No.3 の例)

- 意味の曖昧さがなくコンピュータプログラムを作成しやすい
- 多様な表記法を許容し、人手で容易に記述でき、読みやすい
- 様々な和音の変位の詳細を一定ルールにより表記する応用力を持つ
- 表示や入力に特殊な環境を必要とせず、ASCII テキストデータとして扱える
- 和声学の教程に適合し、かつ世界的に受け入れられやすい

これらの特徴を持つ和声ラベルには、和声学を習得した音楽大学学生によるラベルデータの作成や修正が効率的に行われる、計算機による読み込みが容易であることから統計解析や言語モデルの作成の効率が向上する、等の長所があると考えられる。

KS notation は、上付き文字と下付き文字が無く、転回形は転回の回数をプライム記号 (') の個数によってより直接的に表現されている等の特徴があり、ブレインテキストとして表記できコンピュータ入力が容易である。

3.2 KS-notation の仕様

以上の方針によって策定した和声ラベル表記の仕様を以下に示す。策定に当たっては、様々な既存の和声表記法と、簡潔な楽譜記述に優れた abc 記述言語¹⁵⁾ を参考にした。

3.2.1 和声記法 (A) 単音、複合音

(A1) 単一あるいは複数の音高を “[]” で囲んで単音あるいは複合音を表す。複合音の場合は曖昧さがないように (スペースを挟んで) 併記することで表現できる。

例 { C 上の長三和音 [CEG], Fis 上の長三和音 [+F+A+C'], ...
主和音 [I III V], 属七第三転回三音省略 [IV V II'], ...

(A2) 音高を表現するには、絶対音高はローマ字音名 (“A” から “G” まで)、調の主音からの相対音高はローマ数字、根音からの相対音高はアラビア数字を用いる。但し、ローマ数字は、長音階の場合は大文字、(自然) 短音階の場合は小文字を用いる。アラビア数字は、根音からの完全音程および短音程を基準とする。

例 { 八長調の主音 [C], 嬰へ短調の三音 [A], ...
任意の長調の主音 [I], 上主音 [II], 属音 [V], 導音 [VII], ...
任意の長調の属七の五音省略 [V 3 7]

(A3) 完全音程あるいは短音程を基準として、変位 (半音上昇, 半音下降, 二半音上昇, 二半音下降) はそれぞれ “+”, “-”, “++”, “--” を, アルファベットあるいはローマ数字の前, あるいはアラビア数字の後に付加する。

例 { Fis [+F], As [-A], Gis [+G], ...
下属音の二半音上 [+IV], 導音の半音下 [-VII], ...
長九度の属九 [V 3 5 7 9+]

(A4) オクターブ上の音を表すことが必要な場合は, “” を付加する。ローマ字 (A...G) の場合は小文字あるいは “” の付加によって表しても良い。

例 { オクターブ上の主音 [I], オクターブ上の導音 [VII"], ...
八長調の主和音の第二転回形 [Gce]=[GC'E"], ...

3.2.2 和声記法 (B) 和音

(B1) 長三和音, 短三和音をそれぞれ根音度の大文字ローマ数字および小文字ローマ数字で表す。これは, 後述する三音に関する短音程からの変位・省略を, 3 の後に +, -, ! などと書く代わりに, ローマ数字の大小で表現できるというものである。

例 { 主和音 I=[I III V], 半音下方導音上の長三和音 -VII=[VII- II' IV'], 上主和音 ii=[II IV VI], 下属音上の短三和音 iv=[IV VI- I']

(B2) 完全音程あるいは短音程を基準として, 構成音に変位がある場合は, アラビア数字の後に音程の半音変位を “+” または “-” の個数で示す。また省略は “!” で示す。

例 { 増五度は “I5+”, 長三度は “I”, 長七度は “I7+”, 減七度は “I7-” 三音省略 I3!=[I V], 属七の五音省略 V5!7=[V VII IV']

(B3) 根音に関する変位・省略は, 1 の後に +, -, ! などを書く代わりに, 根音の前にその記号を書くことができる。

例 { 長調の導音の半音下方変位の上の長三和音: -VII = vii1- 属七の根音省略: !V7 = V1!7
II=[III V], !V=[VII II'], +v-=[+V VII II'], ... イ短調の減七 [+G B D' F'] = vii-7- = !V9

(B4) 五音に関する完全五度からの変位・省略は, 5 の後に +, -, ! などを書く代わりに, 5 を省略できる。

例 { 導音上の減三和音: vii- = vii5- 属七の根音省略: !V7 = V1!7 変位は, 音省略は, 以上の前に “!” を付加する。!II=[III V], !V=[VII II'], +v-=[+V VII II'], ...

3.2.3 和声記法 (C) 七度和音, 転回, 調

(C1) 七度の和音, 九度の和音, 十一度の和音などは, それぞれ上記のあとにアラビア数字 “7”, “9”, “11” を付加する。九度の和音の場合, 短七度音の存在を既定とし, 長七度音である場合は 7+ を, 存在しない場合は 7! を 9 の前に付加する。11 度以上の場合も同様。

例 { 属七根音省略 !V7=[VII II' IV']=[VII 3 5-], 主音上の七の和音 I7+=[I III V VII]=[I 3+ 5 7+]

(C2) 絶対調はローマ字音名 (“A” から “G” まで), 相対調はローマ数字にそれぞれ前に変位 (“+”, “-” など) を付加して後に “:” を付加して表す。

例 { 短三度上の調 -III:

(C3) 借用和音は, 相対調と上記の和音記号を結合して表現する。


```

% RWC-MDB-C-2001 No.2
% Mozart: Symphony No.40 in g minor, 1st Movement
@K=g @M=2/2
% Exposition (提示部)
% bars 1-9 (theme I)
||: i | i | i | i | ii-7''' | ii-7''' | V7' | V7' | i |
% bars 10-15
!V9''' | i' | !V9''' | i' | {v: ii-7 | !V-9'' |
% bars 16-19
I! 1/2!V9''' 1/2!V9'''/I | I 1/2!V9''' 1/2!V9''' |
I 1/2!V9''' 1/2!V9''' | I !V9'''/I | !V9'''/I |
% bars 20-24
I} V3!7'' | i V' | i | i | iv7' |
% bars 25-27
iv7'={III: ii7 | V7 | V7 |
% bars 28-33
I | V' | IV' | I'' | ii' | I'={V: IV' |
% bars 34-37
V7 | V7 (!V9''') | V7 (V!9''') | V7 |
% bars 38-43
I} V7 | i'' V: !V9''/I | V V7 |
i'' V: !V9''/I | V z V z | z |
% bars 44-51(theme II)
3I' V: V! | V7 | i!3! !2vi V: V7' | 2I'' 1!V' (1z) |
3IV! vi: V! | ii: V!7 V: V! V!7 I! | ii' V7 | I! |
% bars 52-57
3I' V: V! | V7 | i!3! !2vi V: V7' | 2I'' 1!V' (1z) |
3ii' vi: V7 | ii: V!7 V: V7 V!7 IV: V7} |
% bars 58-65
{-I: 3V7 I!'' | 3!V7 I!'' | 3!V7 I!'' | 3!V7 I!'' |
V7} | {III: V: !V9' | I'' | V7 |
% bars 66-71
I [2D -E | E F F+ G | IV: !V7''' IV! | !V7''' I' | IV! | V |
% bars 72-76
I ( 3I' ) | V7''' | I' | V7''' | I' |
% bars 77-79
vi: V7 | vi ii' | I'' V7 |
% bars 80-84
I | V7' | I | V7' | I |
% bars 85-87
vi: V7 | vi ii' | I'' V7 |
% bars 88-94
I | I V7' | !I! I' IV V7 | I | I V7' |
!I! I' IV V7 | vi IV I'' V7 |
% bars 95-100
I V7 | I V7 | I V7 | I V7 I V7 | I (3z) |
V7'' (3z) :||
%
% Development (展開部)
% bars 101-104
i (z) {+vii: {III: !V9'''' (z) | !V3!7'' |
2I! !V' } V3!7'' | i V' |
% bars 105-111
i | i | 3V: V7'''' ii7'''' | ii-7'''' | V7' | V' } |
{+vi: 3V: !V9'''' iv7'' |
% 112-114
ii-7'''' | V7' | V7 |
% bars 115-122
i'' V7 | i'' (3i') | i iv' | iv' {v: V3!7 | i' V7'' |
i | i iv: V!7' 2iv | iv' (iv) } {IV: V7 |
% bars 123-130
I'' V7 | I'' (3I') | I IV' | IV!={III: V! V3!7'''' |
I' V7'' | I | I! IV7+ | ii'' !V' |
.....

```

図5 KS記法データ例: W. A. Mozart: 交響曲第40番第一楽章 (途中まで)

を行うには必ずしも理想的なバランスではない。

4.2 データ作成作業

和声ラベルデータ作成作業は、三段階に分けて実施された。

第一次作業では、各楽曲全曲の和声分析を、高度な和声学を習得した音楽大学作曲専攻生・卒業生5名と2名の演奏専攻生によるチームにより実施した。一般的に、作曲専攻者の分析からは論理的解釈の特徴が、演奏専攻者の分析からは幼少時に身につけた和声感による解釈の特徴が見受けられることから、解釈に多様性を含ませる目的で、分析者に幅を持たせた。

第二次作業では、和声分析力がありデータ入力に適した音楽大学作曲専攻生と卒業生各1名により、第一次作業で分析された結果をKS notationに変換した。また、それぞれの作業で、専門家による修正が行われた。

第三次作業では、和声の長さの記述によりMIDIデータに変換して和音を演奏することができることから、楽譜を参照しながら和声を聴き、記述誤りの有無をチェックした。この

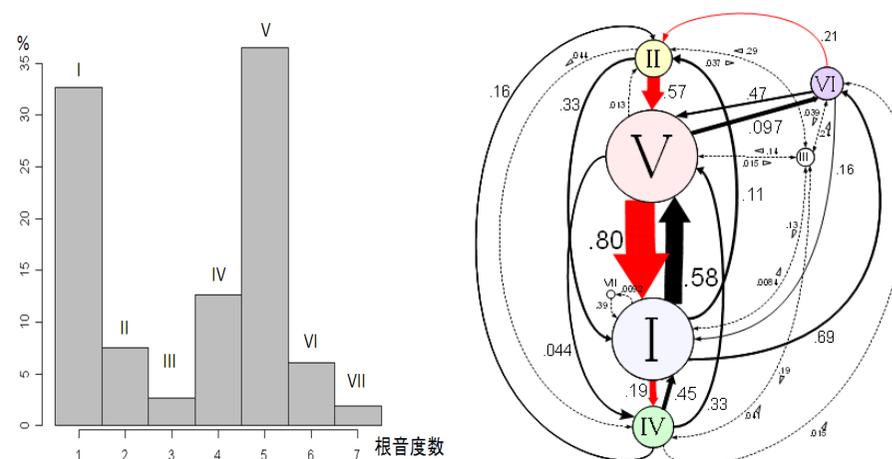


図6 50曲のKSNデータから求めた和音出現(unigram)確率⁸⁾と遷移(bigram)確率。円の面積は和音の生起確率(unigram)を表し、矢印の太さは和音間の遷移確率(bigram)を表す。II, V, Iの間のドミナント進行とIV, I, Vの間のサブドミナント進行の確率が大きい。

機能は、大きなデータセットを作成する場合に極めて有用である。

図5に、KSNデータの例を示す。

5. KSNデータの統計解析と和声モデルへの展望

5.1 和声の種類に関する出現頻度と遷移確率等に関する統計解析

作成したKSNデータに対して統計解析を行った。ラベルデータは、和声系列をなす和声ラベルの合計個数は17764、和声ラベルの種類数は1026、種類の出現頻度から算出したエントロピーは7.08bitであった。

以下、川上ら⁸⁾からの再掲も含めて統計解析とその解釈について述べる。

主音に対する和声の根音の音程のみに着目し、和声の出現頻度とその間の関係を遷移確率を求めた。結果を図6に示す。

ドミナントであるVが最も多く、次にトニックであるI、更に次にサブドミナントであるIV、そしてII, VIの順に出現頻度が高い。7割近くの和声がIとVのみによって占められていた。更に、和声学で示される基本的な進行であるドミナント進行やサブドミナント進行が、主要な和声II, V, I, IVの間で5割以上の高い確率で生起していた。V-Iとなる確率に関しては8割近くであった。これは借用和音の多くがV-Iで形成されていることにも

起因する。I, V に続いて出現頻度の多い IV は、ドミナント、トニックどちらにも進行できるサブドミナント機能の和音で、使用頻度が多いことが予測されたが、統計的にも同様の結果が得られた。また、II はほとんどのケースで進行がドミナントに限定されるものの II V I 進行は機能と声の最も終止感の強い、カデンツの典型的進行である。予測された使用頻度の多さが、統計解析によって実証された。また、楽曲では偶成和音が経過的に置かれる例が多く見られるが、II I, VI V 進行の遷移確率も高くなった。音楽音響信号から和声を推定する際、和声の種類や遷移確率は、明らかに調に依存して偏っているため、調が推定出来れば和声の推定精度も向上すると考えられる。

そのほか、転回形は基本形（特に I と V）が最も多かった。楽曲は和声の連鎖によって節を作るため、その都度の締めくくりとなる終止形では基本形が多く使われることが統計的にも示された。

転調の傾向に関して、属調や下屬調へ転調する確率が 20%弱であり、他の調に比べ頻度が高いこと等がわかった。以上の結果から、実曲であるクラシック音楽 50 曲に関して、和声学から大きく逸脱しているということはないという知見が得られた。

音楽音響信号から和声を推定する際、和声の種類や遷移確率は、明らかに調に依存して偏っているため、調が推定出来れば和声の推定精度も向上すると考えられる。

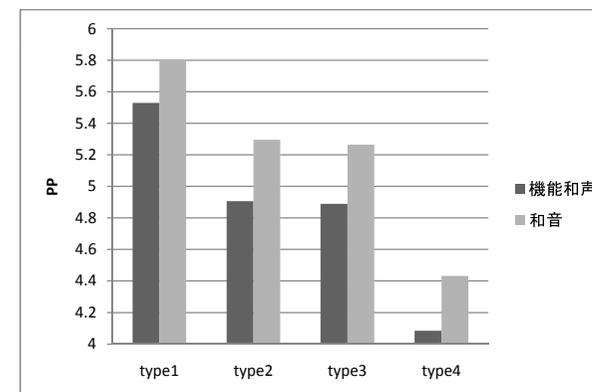
そのほかの統計として、転回形は基本形が最も多く第三転回形が少ないこと、転調の傾向に関して、属調や下屬調へ転調する確率が 20%弱であり、他の調に比べ頻度が高いこと等がわかった。以上の結果から、実曲であるクラシック音楽 60 曲に関して、和声学から大きく逸脱しているということはないという知見が得られた。

5.2 N-gram 和声モデルの perplexity

音響入力からの和声認識は、音声認識と同型の問題である。音声認識は、言語モデルの確率的な拘束の下で音素の尤度を最大にする音素系列を探索する問題と捉えられる。ここでは、言語モデルは極めて大きな役割を担う。和声認識においても、和声がどのような系列をなすのか確率的な拘束は有効に利用することができる^{10),16)}。以下、既報告⁸⁾と重複する部分もあるが、本データベース構築の第一の狙いはここにあるので、再び触れることにする。

和声進行のモデルとして N-gram モデルを検討した。現代の連続音声認識においては、単語間 bigram や trigram が言語モデルとしてよく使われ、効果が実証されている。言語モデルの拘束力を議論する上で、もっともよく用いられるのは perplexity (情報源の持つエントロピーの 2 の冪乗) である。

音楽においては、和音連鎖がどれほどの拘束力を持つのか、それが調と機能と声を考慮したモデルの方が拘束力が強い良いモデルなのかどうか論点である。図右の機能と声モデルは、言語モデルでは、話題別の言語モデルがありそれらが確率的につながれているという構造に喩えることができる。但し、話題によって語彙は変わらないが、同一の語の品詞が話題によって変化し、出現頻度や連鎖頻度も変化する状況と考えることができる。機能と声



- (type1) 和声に関して記述された全ての情報を使用
- (type2) 転回形の情報を無視
- (type3) type2 に対して更に、11 度 13 度音の付加/変位/省略、テンションを無視
- (type4) type3 に対して更に、ベース、根音、5 度音、9 度音の付加/変位/省略を無視

図 7 機能と声モデルとポピュラー和声モデルの perplexities の比較 (再掲⁸⁾)

慮しない図左のような従来のモデルは、話題に分けないモデルに喩えられる。

図 7(再掲⁸⁾) は、和音を細分類から粗分類まで 4 種のまとめ方に関して trigram モデルに基づく言語情報源の testset perplexity を算出した結果である。但し、統計的なスパースネスに対処するために Wintten-Bell 法による平滑化を施してある⁸⁾。この結果から、機能と声に基づく方が和声の言語モデルとして有効であることが示された。音楽学的な見地から言えば、調性に依存した機能と声理論に基づく作曲の方が、和音間にある確率で推移するランダムな作曲より、より古典的な曲らしい曲が作れる、という意味に理解できる。

6. おわりに

本稿では、機能と声解析の記述仕様 (“KS notation”) を策定した。機能と声記述のために、和音、転回、借用和音、省略、変位、転調、付加音などの記述を可能とし、さらに楽譜なしで演奏が可能のように音価も表現した。また、人間とコンピュータ双方の可読性の両立させコンパクトに表現できるようにした。

この仕様に基づいて和声の専門知識を持つ作業員により、人手により機能と声解析を行ってデータを作成した。データ作成には、RWC 音楽データベース所収のクラシック曲 50 曲について、人手により機能と声解析してデータを作成した。

そのデータを統計解析し、音楽的な見聞から説明を試みた。そして、和声の言語モデルとして、機能音声モデルが従来のモデルより工学的音声モデルとして優位であることを perplexity の観点から示した。

今後は、さらに音楽学的な統計解析と同時に、和声系列のモデル化により適した言語モデルについて検討していく方針である。

参 考 文 献

- 1) 川上 隆, 中井 満, 下平 博, 嵯峨山 茂樹, “隠れマルコフモデルを用いた旋律への和声付け,” 平成 11 年電気関係学会北陸支部大会講演論文集, F-61, Oct., 1999.
- 2) 川上 隆, 中井 満, 下平 博, 嵯峨山 茂樹, “隠れマルコフモデルを用いた旋律への自動和声付け,” 情報処理学会研究報告, 99-MUS-34, pp.59-66, Feb., 2000.
- 3) A. Sheh *et al.*, “Chord segmentation and recognition using EM-trained hidden markov models,” *Proc. ISMIR*, pp. 183-189, 2003.
- 4) J. P. Bello and J. Pickens, “A robust mid-level representation for harmonic content in music signal,” in *Proc. ISMIR*, pp. 304-311, 2005.
- 5) K. Lee and M. Slaney, “Acoustic chord transcription and key extraction from audio using key-dependent HMMs trained on synthesized audio,” *IEEE Trans. on Audio Speech and Language Processing*, vol. 16, no. 2, pp. 291-301, 2008.
- 6) 内山 裕貴, 宮本 賢一, 西本 卓也, 小野 順貴, 嵯峨山 茂樹, “調波音を強調したクロマに基づく音楽音響信号からの自動和音認識,” 日本音響学会春季研究発表会講演集, pp.901-902, Mar., 2008.
- 7) 内山 裕貴, 宮本 賢一, 西本 卓也, 小野 順貴, 嵯峨山 茂樹, “調波音・打楽器音分離手法を用いた音楽音響信号からの自動和音認識,” 情報処理学会研究報告, 2008-MUS-76, 23, pp.137-142, Aug., 2008.
- 8) 川上 大輔, 金子 仁美, 嵯峨山 茂樹, “和声ラベルデータの作成と和声進行の統計解析,” 情報処理学会研究報告, Feb., 2010.
- 9) 川上 大輔, 金子 仁美, 嵯峨山 茂樹, “機能音声ラベルデータの作成と統計解析,” 日本音響学会春季研究発表会講演集, pp.979-980, Mar., 2010.
- 10) 上田 雄, 小野 順貴, 嵯峨山 茂樹, “調波音/打楽器音分離手法とチューニング補正手法を用いた音楽音響信号からの自動和音認識,” 情報処理研究会報告, 2009-MUS-81, pp.1-6, 2009.
- 11) Yushi Ueda, Yuuki Uchiyama, Takuya Nishimoto, Nobutaka Ono, Shigeki Sagayama, “HMM-Based Approach for Automatic Chord Detection Using Refined Acoustic Features,” *Proc. of ICASSP*, Mar., 2010.
- 12) C. Harte *et al.*, “Symbolic representation of musical chords: A proposed syntax for text annotations,” *Proc. ISMIR*, pp. 66-71, 2005.
- 13) M. Goto *et al.*, “RWC music database: Popular, classical, and jazz music databases,” *Proc. ISMIR*, pp. 287-288, 2002.
- 14) 島岡譲他: 和声 理論と実習 I, II, III, 音楽之友社, 1964, 1965, 1967.
- 15) abc 言語 <http://abcnotation.com/>
- 16) 諸岡孟, “確率的音楽生成モデルに基づく自動和声解析の研究,” 東京大学情報理工学系研究科修士論文, Mar. 2008.