

# 和音認知に関する心理物理モデル

藤澤 隆史<sup>1</sup>, Norman D. Cook<sup>2</sup>, 長田 典子<sup>1</sup>, 片寄 晴弘<sup>1</sup>

<sup>1</sup> 関西学院大学 理工学研究科／ヒューマンメディア研究センター

<sup>2</sup> 関西大学 総合情報学部

## 概要

和音／和声 (chord/harmony) は、メロディ (melody)、リズム (rhythm) とともに音楽を形作る重要な要素である。音楽の物理的な音響的特徴とその心理的な印象や感性との関連性について定量的に評価するために、本研究では和音性についての評価モデルを構築した。和音性は、(1) 協和性 (澄んだー濁った)、(2) 緊張感 (緊張したー落ちついた)、さらに長調か短調かといった性質を決定する(3)モダリティ (明るいー暗い、うれしいー悲しい) から構成される。本研究において提案されたモデルと、これまで経験的に知られている様々な和音タイプや、得られた実験データとの整合性を確認し、モデルの妥当性について検証した。

## A Psychophysical Model of Chord Perception

Takashi X. Fujisawa<sup>1</sup>, Norman D. Cook<sup>2</sup>, Noriko Nagata<sup>1</sup> and Haruhiro Katayose<sup>1</sup>

<sup>1</sup> Research Center for Human Media, Kwansai Gakuin University

<sup>2</sup> Department of Informatics, Kansai University

## Abstract

A psychophysical model designed to explain the phenomena of resolved/unresolved harmonies and the major/minor modalities in traditional Western diatonic music is presented. The model uses solely the acoustical features of the pitch combinations for calculation of the total “dissonance”, “tension” and “modality” of chords. Dissonance is defined as a 2-tone effect, similar to the model of Plomp & Levelt. Tension is defined as a 3-tone effect due to the relative size of intervals, following the idea of “intervallic equivalence” by Leonard Meyer. The total sonority of any number of tonal combinations can be computed on the basis of these two concepts.

### 1. はじめに

音楽情報処理において、楽音から受けとる感性情報を定量的に評価することは重要なプロセスである。メロディやリズムとともに、和音／和声 (chord / harmony) は音楽の構成を担う重要な要素であるとされており、例えば、長

調と短調の和音から、それぞれ“明るさ”や“暗さ”といった視覚的印象、“喜び”や“悲しみ”など気分・感情に関連した印象[1]を受けるように、和音／和声が豊かな感性情報を含んでいることを指摘する研究は数多い。また学童期以前の子どもであっても、和音やメロディにおい

て調の長短を有意に識別するといった知見 [2][3]も考慮すれば、和音/和声という音楽的要素は人間生活において深く根ざしているものであるとすることができる。なぜ長調の和音やメロディが楽しく聴こえ、反対に短調は寂しく聴こえるのか、さらにある和音はその他の和音に比べてどの程度楽しさの印象を伴って聴こえるのかということについて十分な解明がなされているとは言いがたい。そこで本研究はそれらの諸問題を解決するために、“和音性”についての定量的評価モデルを構築することを試みた。

## 2. これまでの和音知覚モデル (協和性)

### 2.1 倍音と協和性

和音/和声に関わるこれまでの先行研究において、その心理物理の対応関係がほぼ明らかにされているのは、音の“協和性”についてである。協和性は主に2音のピッチ距離(音程)から決定される性質である。ある2つの純音において、そのピッチ間の距離が小さい(1-2半音程度)場合、それらの音は知覚的に明確に2音として分離ができないので濁って聴こえるが、逆に距離が大きい場合、その濁りは消失するという仮定に基づいてモデルが構成されている[4] (Fig.1の $F_0$ )。

和音の構成音が複合音である場合、その協和性(不協和度)はそれぞれ倍音の影響を受ける。例えばCとF#の場合、その $F_0$ 間の音程は6半音であるから不協和度の値は低いものとなるが、Cの $F_2$ (G)とF#の $F_1$ (F#)の音程が1半音であることから、総合すると不協和度の値は高くなる (Fig.1の $F_0-F_2$ )。

### 2.2 不協和度曲線

ある構成音の不協和度を決定する式は、これまでの先行研究で用いられたものとはほぼ同様のものであり、モデルから得られる不協和度の値のパターンはほぼ一致している[5][6]。ある2つの純音の周波数 $f_1, f_2$ から構成される音程を $x_{12}$ とし、それぞれの音量を $v_1, v_2$ とした場合、

その不協和度( $d$ )は以下の式によって定義される。

$$d = v_{12} \alpha_3 \left[ \exp(-\alpha_1 x_{12}^\beta) - \exp(-\alpha_2 x_{12}^\beta) \right] \quad (1)$$

ここで、 $v_{12}$ は2音の音量 $v_1, v_2$ から定義される値であり、本研究では先行研究[67]に従い $v_1, v_2$ の積とした。

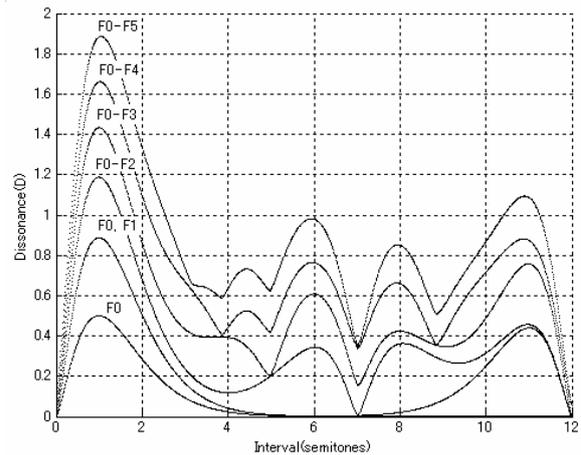


Fig.1 不協和度( $D$ )の理論曲線

さらに構成音が複合音である場合、音1の $i$ 倍音の周波数を $f_{1i}$ と音2の $j$ 倍音の周波数を $f_{2j}$ から構成される音程を $x_{12}$ とし、またそれぞれの音量を $v_{1i}, v_{2j}$ とした場合、その不協和度( $D$ )は以下のように定式化できる。

$$D = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} d(x_{ij}, v_{ij}) \quad (2)$$

本研究では、3節以降は特に3和音について議論するが、3和音の場合、含まれる音程の数も3つとなる。3つの音程からそれぞれの不協和度( $D$ )が計算されるので、平均をもってその3和音についての不協和度であると定義した。

## 3. 和音性モデル (緊張度・モダリティ)

### 3.1 協和性と和音性

長調や短調など各種の和音がもつ印象や感性は、協和性理論に基づいた音の音響学的特性

から説明可能だとする立場は存在するが[7][8], その予測と経験もしくは実験データ[9]との整合性があるとは言いがたい (Table.1 を参照). “和音性”という語を“われわれが和音から受けとる感性的性質の総体”と定義するならば, この不整合は, 和音性の全てを協和性のみから説明することはできないことから生じるものである[11][12].

そこで本研究では, Mayer の古典的知見[10]に基づくことで, 和音に関する知覚的性質 (= 和音性) を新たに 2 つ仮定した. それらは“音程の等しさ (intervallic equivalence)”に基づくもので, 例えば, ある 3 つの純音がそれぞれ等しい音程で離れている場合, その 3 和音は“緊張的”で“未解決的”な響きをもつと仮定する. 反対にそれぞれの音程が等しくない場合には, “落ち着いた”もしくは“解決的”な響きをもつ (Fig. 2). この和音性を“緊張性”と定義する.

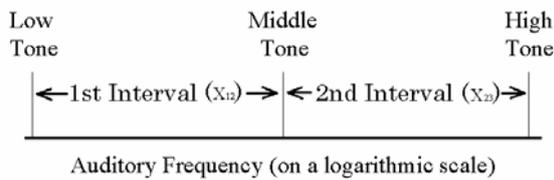


Fig.2 第 1 音程 ( $x_{12}$ ) と第 2 音程 ( $x_{23}$ )

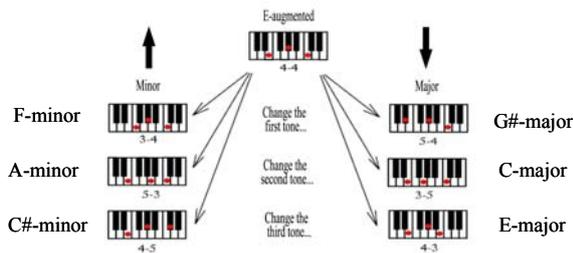


Fig.3 解決的和音と長調/短調の関連性[11]

さらに“解決的”な響きには, 長調的に明るさや楽しさの印象を伴う場合と, 短調的に暗さや寂しさの印象を伴う場合に 2 通りがあり, それぞれも同様に 3 音から構成される相対的な音程構造から定義されると仮定した. 本研究で

はこれを“モダリティ”と定義する. “緊張性”との関連から調の長短についての概念を捉えなおすことで, 従来の音楽理論に基づかない, 定量的な和音性評価モデルを構築することが可能となった.

### 3.2 緊張度の定義

ある 3 つの純音について, それぞれの周波数を  $f_1, f_2, f_3$  ( $f_1 < f_2 < f_3$ ) とした場合,  $f_1$  と  $f_2$  から定義される音程を  $x_{12}$ ,  $f_2$  と  $f_3$  から定義される音程を  $x_{23}$  とし, またそれぞれの音量を  $v_1, v_2, v_3$  とした場合, 緊張度( $t$ )以下のように定式化される.

$$t = v_{123} \exp \left[ - \left( \frac{x_{23} - x_{12}}{\gamma} \right)^2 \right] \quad (3)$$

ここで,  $v_{123}$  は 2 音の音量  $v_1, v_2, v_3$  の積であり,  $\gamma$  は定数で 0.60 である. 式 (3) によって定義される緊張度 ( $t$ ) の基本曲線を Fig.4 に示す (ただし,  $v_1 = v_2 = v_3$ ).

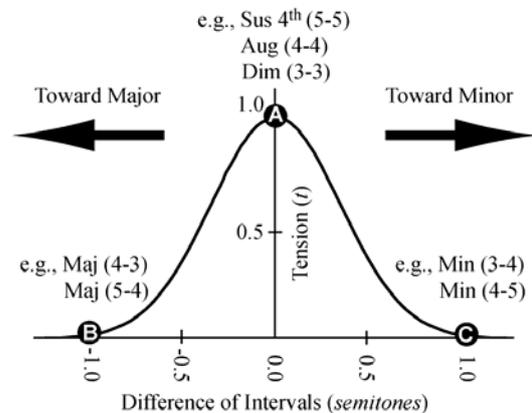


Fig.4 緊張度( $t$ )の基本理論曲線

3 和音の構成音が倍音成分を含む複合音である場合の緊張度 ( $T$ ) は, 音 1 の  $i$  次周波数を  $f_{1i}$ , 音 2 の  $j$  次周波数を  $f_{2j}$ , 音 3 の  $k$  次周波数を  $f_{3k}$  とし, またそれぞれの音量を  $v_{1i}, v_{2j}, v_{3k}$  とした場合, 以下のように定式化できる.

$$T = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} \sum_{k=0}^{n-1} t(x_{ij}, y_{jk}, v_{ijk}) \quad (4)$$

式(4)によって定義された緊張度 ( $T$ ) の理論曲線を, 第 1 音程が 3 半音 ( $x_{12}=3.0$ ) である場合を例として Fig.5 に示す.

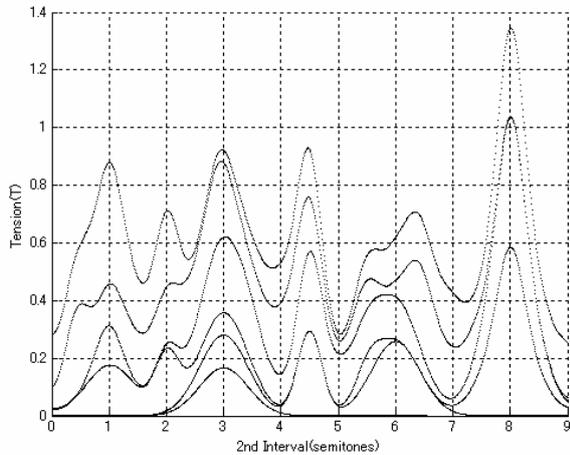


Fig.5 緊張度( $T$ )の理論曲線

### 3.3 モダリティの定義

モダリティはある 3 和音が長調的か短調的かを示す指標値であり, 値が正である場合は長調的, 負である場合は短調的となる. 緊張度( $t$ )と同様に, ある 3 つの純音の周波数  $f_1, f_2, f_3$  ( $f_1 < f_2 < f_3$ ) から構成されるそれぞれの音程を  $x_{12}, x_{23}$  とした場合, モダリティ( $m$ )は以下のように定式化される.

$$m = -v_{123} \left[ \frac{2(x_{23} - x_{12})}{\varepsilon} \right] \exp \left\{ - \left[ \frac{-(x_{23} - x_{12})^4}{4} \right] \right\} \quad (5)$$

$e$  は定数で  $e = 1.214$  である. 式 (5) によって定義されるモダリティ( $m$ )の基本曲線を Fig.6 に示す (ただし,  $v_1 = v_2 = v_3$ ). 構成音が倍音成分を含む複合音である場合のモダリティ ( $M$ ) は, 緊張度( $T$ )の定義時と同じ要領で以下のように定式化される.

$$M = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} \sum_{k=0}^{n-1} m(x_{ij}, y_{jk}, v_{123}) \quad (6)$$

また式(6)によって定義されたモダリティ ( $M$ ) の値を Fig.7 に示す ( $x_{12}=3.0$  の場合).

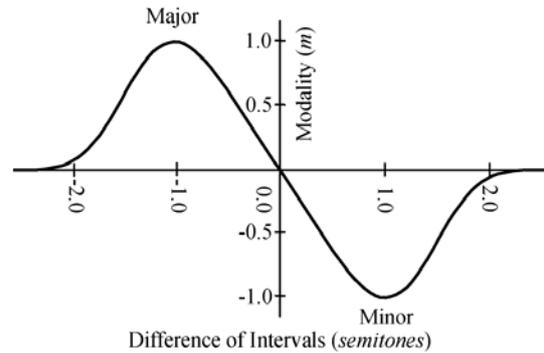


Fig.6 モダリティ( $m$ )の基本理論曲線

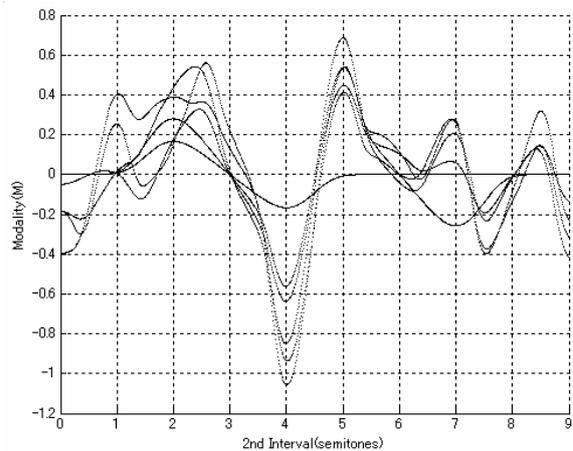


Fig.7 モダリティ( $M$ )の理論曲線

### 3.4 不安定性の定義

不安定性はある 3 和音がどれほど “不安定” な印象で知覚されるかについての指標である. 不安定度 ( $I$ ) は不協和度( $D$ )と緊張度( $T$ )から算出され, 以下の式によって定式化される.

$$I = \frac{1}{3} \sum_{i=1}^3 D_i + \delta T \quad (7)$$

ここで,  $\delta = 0.207$  であり, 不安定度 ( $I$ ) に対する緊張度 ( $T$ ) の相対的影響度を決定するものである.

## 4. 結果と考察

### 4.1 和音の不安定性についての検討

まず和音の不安定性について, 先行研究と本研究のモデルについて比較を行なう. Table.1 に示されたように, 主な 3 和音を聴取した実験データでは[11], musician か non-musician か

によって判断に多少の違いがあるが、一般的に長調(Maj.) > 短調(Min.) > 減和音(Dim) > 増和音(Aug.)のカテゴリー順で“安定的”だと評価される。先行研究における協和性に基づいた3和音の安定性についての予測値はその評価と矛盾している(下線部)のに対し、本研究で提案されたモデルの値はより整合的である。

Table.1 不安定性の先行研究との比較

Chord Class	Interval Struct.	Exp. Roberts	Theory				
			P&L	K&K	Parncutt	Sethares	
Major	4-3	I	4	1	1	4	1
	3-5		<u>11</u>	<u>11</u>	6	<u>8</u>	5
Minor	5-4	II	2	6	3	2	4
	3-4		4	1	4	4	2
Dim.	4-5	III	2	6	6	2	3
	5-3		<u>11</u>	<u>11</u>	<u>10</u>	<u>8</u>	6
Aug.	3-3	IV	13	6	9	12	12
	3-6		8	9	5	10	7
Sus. 4th	6-3	IV	8	9	8	10	10
	4-4		10	13	2	12	13
Sus. 4th	5-2	IV	<u>6</u>	<u>1</u>		<u>6</u>	8
	2-5		<u>6</u>	<u>1</u>		<u>6</u>	11
Sus. 4th	5-5	IV	1	1		1	<u>9</u>

#### 4.2 主な3和音についての検討

主な3和音における、それぞれの指標値について Table.2 に示す。

Table.2 主な3和音の理論値

	Interval	Exp.	Theory				Rank
			D	T	I	M	
Maj	4-3	Root	0.504	0.583	0.624	3.780	1
	3-5	1st Inv.	0.641	0.836	0.814	2.611	5
	5-4	2nd Inv.	0.498	1.366	0.780	3.825	4
Min	3-4	Root	0.504	1.158	0.744	-4.209	2
	4-5	1st Inv.	0.498	1.246	0.756	-3.075	3
	5-3	2nd Inv.	0.641	0.950	0.838	-2.763	6
Dim	3-3	Root	0.764	3.223	1.431	0.361	12
	3-6	1st Inv.	0.695	2.026	1.114	0.176	7
	6-3	2nd Inv.	0.695	2.420	1.196	-0.895	10
Aug	4-4	Root	0.611	6.701	1.998	0.384	13
Sus4	5-2	Root	0.715	2.226	1.175	-0.001	8
	2-5	1st Inv.	0.715	2.438	1.219	-0.061	11
	5-5	2nd Inv.	0.569	3.005	1.191	-0.304	9

まず前節で検討したように、和音の不安定性では増和音が最も不安定であり、長調の和音が最も安定的であると結果となった。次に、長調的か短調的かという性質を示すモダリティ(M)の値について検討してみると、まず長調のカテゴリーに当てはまる和音は+3 前後の値、逆に

短調のものは-3 前後の値となり、それぞれ正しく評価されていることが分かる。さらに増和音、減和音のモダリティの値は0 前後の値であり、長調・短調の和音ほど楽しさや寂しさ、明るさや暗さといった印象を伴わないという評価の値となった。

#### 4.3 和音の推移と理論値

和音の推移によって、和音性(緊張感・モダリティ)の値がどのように変化するかを検討するために、2つの例を Fig.8 に示した。Fig.8(a)では(Csus4→C)、緊張感の減少とモダリティの増加が見られることから、明るさの伴う終止感が得られること、また Fig.8(b)では(Caug→Am)、緊張感の減少とモダリティの減少が見られることから、暗さの伴う終止感が得られることを示唆している。

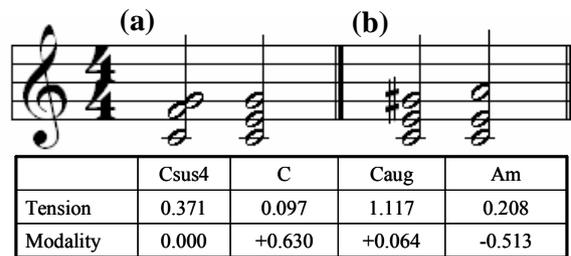


Fig. 8 和音推移と緊張度・モダリティの変化。

#### 4.3 実験結果との比較

モダリティにおけるモデルの予測値と実験から得られたデータとともに Fig.9 に示す。

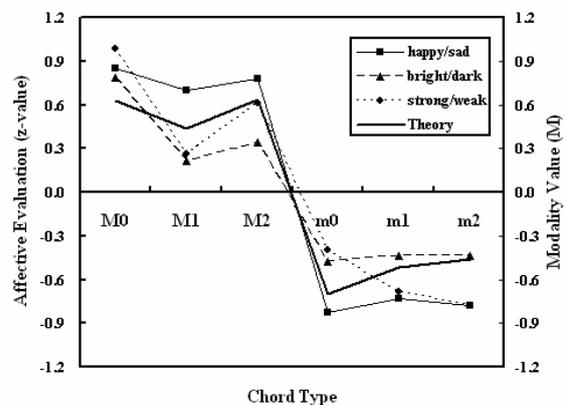


Fig.9 理論値と実験結果の比較

実験参加者はそれぞれ順に 18, 20, 66 人の大学生であり, 全員が 8 年以下の音楽経験しかもたない。モダリティに関して和音を評価する形容詞対はそれぞれ, うれしいー悲しい, 明るいー暗い, 強いー弱い が用いられた。転回型を含む 6 種類の長短調の和音を様々なピッチ水準で聴取させ, 5 段階尺度で評定してもらった。

本研究の予測値 (太線) は実験値とよく適合するものであった。まず長調と短調を区別するその相対的評価は理論値と実験値間で非常に共通していた。さらに興味深い結果として, モデルから長調の第 1 転回型は他の長調の和音に比較してモダリティの値が低いと予測されるが, 3 つの実験それぞれにおいてその傾向が確認された。

## 5. おわりに

本研究では, 和音を定量的に評価するためのモデルとして, 音程構造に基づいた“緊張性”および“モダリティ”という新たな構造指標を提案し, 先行研究との対応を含めその妥当性について検討した。本研究のモデルは和音という側面から, 音楽という複雑な心理物理現象を明らかにするための基盤となりうるものである。従来の音楽理論では, 和音や和声については定性的にしか把握されてこなかったが, 定量化を行うことで音楽情報処理のための新たな特徴量として用いることが可能である。例えば, 緊張性の指標を用いることで楽音の緊張ー弛緩構造と和声進行を対応づけることも考えられる。また, 他の心理物理データや脳科学との対応を明らかにすることもできよう。さらに本モデルは, 音響的特徴からわれわれが直接的に受ける印象や感性をモデル化したものあることから, 他の音響現象, 例えば, 音声の音楽性について評価する [15] といった応用可能性についても考えられる。

## 参考文献

- [1] Juslin, P.N. & Sloboda J.A. (edt.) (2001). *Music and Emotion: Theory and research*. Oxford: Oxford University Press.
- [2] Trehub, S.E., Morrongiello, B.A., & Thorpe, L.A. (1984). Children's perception of familiar melodies: the role of intervals, contour, and key. *Psychomusicology*, **5**, 39-48.
- [3] Kastner, M.P., & Crowder, R.G. (1990). Perception of major/minor: IV. Emotional connotations in young children. *Music Perception*, **8**, 189-202.
- [4] Plomp, R. & Levelt, W.J.M. (1965). *Journal of the Acoustical Society of America*, **38**, 548-560.
- [5] Kameoka, A., & Kuriyagawa, M. (1969). Consonance theory: Parts I and II. *Journal of the Acoustical Society of America*, **45**, 1451-1469.
- [6] Sethares, W.A. (1999). *Tuning, Timbre, Spectrum, Scale*, Springer, New York.
- [7] Terhardt, E. (1974). Pitch, consonance and harmony. *Journal of the Acoustical Society of America*, **55**, 1061-1069.
- [8] Parncutt, R. (1989). *Harmony: A psychoacoustical approach*, Springer, Berlin.
- [9] Roberts, L. (1986). Consonant judgments of musical chords by musicians and untrained listeners. *Acustica*, **62**, 163-171.
- [10] Meyer, L. (1956). *Emotion and Meaning in Music*, Chicago University Press, Chicago.
- [11] Cook, N.D. (2002). *Tone of Voice and Mind*, John Benjamins, Amsterdam.
- [12] Cook, N.D. & Fujisawa, T.X. (2006). The Psychophysics of Hermony Perception: Harmony is a Three-Tone Phenomenon. *Empirical Musicology Review*, **1**(2), 106-126.
- [13] Cook, N.D. Fujisawa, T.X. & Takami, K. (2006). Evaluation of the Affective Valence of Normal Speech Using Pitch Substructure. *IEEE Transactions on Audio, Speech, and Language Processing*, **14**(1), 142-151.