

聴覚モデルを用いた楽音の協和／不協和に関する考察

武田 晃[†]

A Study of Musical Consonance / Dissonance Using an Auditory Model

Akira Takeda

1. はじめに

ヒトの聴覚は、音声によるコミュニケーションという高度な情報処理を行い、その一方で音楽を感じることができる。音楽はヒトに心理的な感動をもたらすが、他の動物にはない特徴である。一体ヒトはなぜ音楽に対して様々な感情を抱くのだろうか。音は物理現象であるがそれが音楽となってヒトの聴覚を通したとき心理的な効果が生じるのはどのような仕組みになっているのか、聴覚モデルシミュレーションを通して探ってみたのが本研究である。

2. 序論

2.1 物理現象としての音

どんなに複雑な音であっても正弦波を合成することで表現できる。音の構成要素は周波数、位相と振幅である。一つの正弦波だけで表現できる音を純音(pure tone)と言う。周波数が近い2つの音を重ね合わせると「うなり(beat)」と呼ばれる緩やかな周波数の音を感じる。加法定理で表現すると、2つのsin波を加算した場合、sinとcosの積で表すことができ、cos項は変調波で「うなり」(包絡線)になり、sin項は搬送波の関係になる。Helmholtzは、2つの音の周波数差によって生じるbeatとroughnessによって不協和感を知覚すると考えた[1]。Setharesは2音の周波数差が小さいとbeatと呼ばれる「うなり」が聞こえ、その周波数差が大きくなるにつれざらざらした音に感じるとroughnessであると説明している[2]。実験で得られた不協和度とbeatとroughnessの関係が図1である。

2.2 音楽を奏でる楽音

1) 楽音

周波数が異なる複数の純音が合成された音を複合音(complex tone)と呼ぶ。特に複合音の中で弦楽器や管楽器のように音の高さ(pitch)が明確な楽器音は周期的波形を持ち、これらを楽音(musical tone)と呼ぶ。周波数が f , $2f$, $3f$, ...の正弦波を合成した音を調波複合音(harmonic

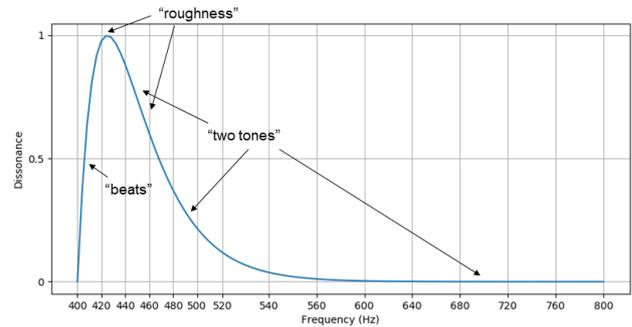


図1 beatとroughnessの関係

complex tone)と呼び、 f を基本周波数(fundamental frequency), 周波数 $2f$, $3f$, ...の各成分を高調波(harmonic)という[3]。

2) 音程(dyad)

2つの音の隔たりのことを音程という。音楽では「度」という言葉で表現する。周波数比が単純な音程ほど響きが美しい。

3) 三和音(triad)[4]

三和音は「主音」に3度上の音と5度上の音を積み重ねることで得られる。和音の基礎となる「根音」がC(ド)の場合、3度上の音が長3度のE(ミ)、かつ「根音」と5度上の音が完全5度をなすG(ソ)の和音を「長三和音」(major triad)と呼ぶ。「根音」のCと短3度のE \flat と完全5度のGの和音を「短三和音」(minor triad), 「根音」のCと短3度のE \flat と減5度のG \flat の和音を「減三和音」(diminished triad), 「根音」のCと長3度のEと増5度のG \sharp の和音を「増三和音」(augmented triad)と呼ぶ。

2.3 ヒトの聴覚器

ヒトの聴覚器官で内耳部分にある蝸牛は、内部がリンパ液で満たされており、その振動は蝸牛内部の基底膜を振動させる。基底膜は厚さと幅が一定でなく、周波数に応じて振動する位置が異なるという特徴がある。さらに、基底膜の振動は感覚受容細胞である有毛細胞を揺らし、神経インパルスを生じさせる。「音の知覚に関する原情報は基本的

[†] 2018年度修士(情報学プログラム), 現所属: 放送大学教養学部

にはすべて基底膜振動に含まれると考えられる[5]」ため、基底膜振動の解析はヒトの聴覚を理解する上で重要である。

2.4 聴覚のしくみ

1) 周波数分析器としての聴覚器

聴覚器は音の高さ（ピッチ）を知覚する分析器の一種と考えられている。フーリエ解析することで音を構成している各正弦波の周波数を求めることに似ている。

2) 音の高さを決めているもの

音には大きさ（loudness）、音の高さ（pitch）、音色（timber）と呼ばれる3つの心理的属性がある[5]。しかし、私たちが耳にする音や楽音は様々な周波数を含んだ複合音であることがほとんどで、単純に音の高さを決めることはできず複合音として音の高さを知覚している。

聴覚において音の高さが何によって決まるか、時間説と周波数説（場所説）があり古くから議論されてきた。時間説は、「うなり」が音を構成している各周波数でなく変調波の周波数を知覚していることをうまく説明できる。周波数説は周波数に応じて基底膜の振動する場所が決まっていることに由来する。

2.5 楽音における協和度／不協和度

純音を聴いただけの時は、その音に対して良し悪しを感じることはあまりない。しかし、「うなり」の例のように2つの音を同時に聴いたときの条件、特に周波数によって良し悪しを感じるようになる。2つの音高の異なる合成音（dyad）の協和度／不協和度に関する研究が多数あり、数理モデルや評価法が提案されている。さらにこれを音高の異なる3つの合成音（triad）の協和度／不協和度に拡張することで、楽音に対する知覚に関する研究も試みられている。

3. 先行研究

3.1 心理的研究に基づく物理モデル

1) 振幅周波数特性と帯域幅

心理的な手法で得られた結果をもとに基底膜の物理モデルとして聴覚フィルタを想定し、聴覚特性について推測することが古くから行われている。ある中心周波数を持つ複数並んだバンドパスフィルタを用いて、心理物理実験からフィルタ形状を予想し基底膜振動を推測するものである。

聴覚フィルタを線形と仮定し、フィルタの帯域幅から振幅周波数特性の推定が試みられ、「臨界帯域（critical band）」の測定を初めて行ったのがFletcher[6]である。Moorは臨界帯域において中心周波数と帯域幅の間に関係があることを示し定式化した[7]。

2) フィルタ形状と音圧依存 [8]

基底膜振動の特性は音圧によって変化する。聴覚フィルタの形状もレベル依存性があり、入力音圧が低い場合にフ

ィルタ形状は鋭く、音圧上昇に伴ってなだらかになる。また、音圧が上昇するとフィルタ利得が減少する。

3) 圧縮特性 [8]

基底膜振動の特性に、入力音圧が大きくなっても振幅が制限される「圧縮特性」がある。ダイナミックレンジの大きい入力音を比較的狭い範囲の振動に変換する。

4) 聴覚フィルタの定式化

ノッチ雑音マスキング実験の結果による聴覚フィルタの推定に、Pattersonはroex(rouned exponential)をフィルタ関数の近似値として導入した[9]。roex(ローエックス)フィルタは片側の周波数領域のみで定義される重み付け関数で、聴覚抹消系の時間応答を表現できないが、実験的にも比較的安定してフィルタ形状を推定できたため広く利用されてきた。時間応答も考慮したフィルタには、線形のガンマトーンフィルタがある。

Meddisらは2つのガンマトーンフィルタを並列に組んだDRNL(Dual-Resonance, Non-Linear)フィルタを導入し圧縮特性や抑圧特性を説明している。また、異なる最適性の観点からIrio and Pattersonはガンマトーンを拡張したガンマチャープを提案した。また、「圧縮型」ガンマチャープ(cGC, compressive gammachirp)や、レベル依存の形状と圧縮特性を一度に導入、さらに、時間的に変化する音に対して動的に特性を変化させる動的圧縮型ガンマチャープ(dcGC, dynamic compressive gammachirp)も提案している。

5) 基底膜振動を表現する聴覚フィルタモデル [10]

基底膜振動のモデルは、その先の神経インパルスへ変換や信号処理は含まず、楽音の協和度／不協和度に直接は結びつき難い。MeddisらはDSAM(The Development System for Auditory Modeling)[11]を、PattersonらはAIM(Auditory Image Module)[12]を提供している。しかし、音声分析には利用できるが、楽音解析としての対応はない。

3.2 不協和度（Setharesモデル）

1) 音程(dyad)の不協和度

心理実験によって得られた不協和曲線をSethares[13]は数式化した。2音の純音の複合音からなる複合音の協和度／不協和度をよく予測している。

2) 加法モデルを三和音の予測に適用

加法モデルとは3音以上の複合音の不協和度をdyadに分解し、不協和度の和を求める。Plomp and Leventによるモデル(P&Lモデル)以降、Kameoka and KuriyamaやSetharesなど複数の研究者が様々なモデルを提案しており、それらのモデルを用いた三和音の予測も行われている。

山本[14]らは、加法モデルは基本周波数とその倍音のみで構成されている静的(static)な音程の不協和度を推定することができるが、楽音に適用しても同様のことができるという報告はないと述べている。その理由に楽音の特徴は振幅変動だけでなく、非調和成分を多く含むことを挙げ、不協和知覚において非調和成分が加法モデルの仮定しない影響を与える可能性があるとしている。

3.3 聴覚モデル IPEM Toolbox [15][16]

IPEM Toolboxは、Matlabでプログラミングされた聴覚モデルによる音楽分析システムである。Auditory Peripheral Module(APM)は、Van Immerseel and Martens [17]によって開発され、音楽知覚にとって重要な機能を実装している。wavファイル等による音データに対して、和音および調性、ピッチ、不協和感、拍子の抽出、音色特徴などの知覚に対処する機能が提供されている。

Auditory Peripheral ModuleとRoughness Moduleの概要を図2に示す。中心周波数の異なる40個並列に並べたバンドパスフィルタ (BPF) のチャンネル(channel)の信号は、神経インパルスに変換され、Roughness Moduleの入力となる。Roughness Moduleを通してchannelの関数(周波数)またはframe内周波数 (Bin No.) の関数としてroughnessのスペクトルが得られるが、最終的に総和をroughnessとしているためスペクトルの特徴が考慮されていない。

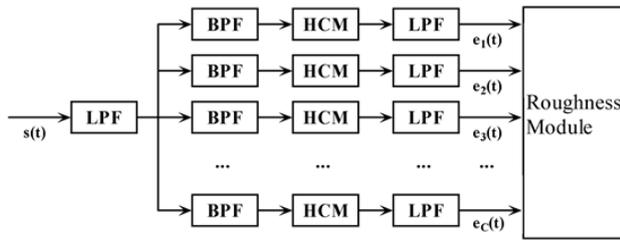


図2 Auditory Peripheral Module と Roughness Module

3.4 調波構造の類似性

Gill and Purves [18]は、世界で広く使用されている音階は、それを構成している音程の音のスペクトルが調波系列とよく類似しているものであることを示した。類似性の測定方法については、2つの音が持つ調波系列の最大公約数で定義される完全調波系列(complete harmonic series)の数を分母に、2つの音が持つ高調波周波数の数を分子としてパーセンテージ (percentage similarity) で表現する。

音程(dyad)の2つの音高だけでなく調波系列とパーチャルピッチも考慮することで、三和音(Maj, Min, Dim, Aug)のランキング予測は良い結果が得られる。

3.5 他の様々な提案やモデル

1) ピッチ顕著性(salience)

Bidelman and Heinz [19]は、聴覚モデルを使って正常および聴覚障害についてシミュレートすることで音程に関する協和と不協和の応答を研究した。彼らはdyadとtriadを分析するために聴覚モデルにauditory-nerve (AN)を使用し、そこからシミュレートされた神経信号を入力にbeat (roughness) とピッチ顕著性(salience)の尺度計算を行った。その結果ANレベルでは、不協和音程/和音に対して協和音程/和音がより強い神経ピッチ(neural pitch)の顕著性を引き出した。協和/不協和評価に、roughnessではなく、刺激のピッチ顕著性を尺度に相関を示した。

2) 二重プロセス理論

Johnson-Laird [20]は、二重プロセス理論を提案している。不協和の心理音響理論はヘルムホルツ(Helmholtz)に従い、基底膜の同領域の周波数が非常に近い部分(基本周波数または倍音)に影響を及ぼし、それとは対照的に、西洋音楽の調性理論は不協和音を高調波原理に反するものと見なし、調性の原理の中にroughnessを埋め込んだ二重プロセス理論(dual-process theory)を提案している。

3) Mc Dermott [21]の非調和による不協和

不協和音の和音は常に大きなbeatを生成するとは限らず、2つの音によって生成されるbeatは音のスペクトルに依存する。それゆえに、楽器間でかなり異なることがある。対照的に、高調波周波数関係の知覚は抽出された周波数振幅にほとんど依存しないためより不変的に音楽構造に関連する可能性がある。彼らは、不協和は非調和周波数を生成する音符の組み合わせに起因することを示唆しているとの知見を示している。

4. 実験と考察

4.1 実験の目的

Bin No.側からroughnessスペクトル構造を観察し、三和音のランキングを並べることで類似性の存在を調べる。

4.2 楽音データ

音色の種類は純音(Pure Tone)、複合音(Complex Tone)、Piano, Organの4種類のwavデータを用意した。

音程は下の音をC4(262(Hz))として、上の音をそれぞれD \flat 4, D4, E \flat 4, E4, F4, G \flat 4, G4, A \flat 4, A4, B \flat 4, B4, C5を組合せた12種類のデータを用意した。

三和音(Chord)はC4(262(Hz))を根音(root)とするMajor(Maj), Minor(Min), Diminished(Dim), Augmented(Aug)の4種類のデータを用意した。

4.3 実験のための予備検討

IPEM ToolboxはChannel側とBin No.側からのスペクトル観察が可能である。Bin No.側からroughnessスペクトルの構造を見ると、三和音のroughnessスペクトル(Maj, Min, Dim, Aug)でそれぞれ異なる位置に主なピークが出現する。このピークの特徴だけからは三和音のランキングは並ばないが、先行研究にもあるようにroughness一つの効果だけでなく、他の効果と合わせて協和度/不協和度を考えてみることにした。

4.4 percentage similarityの検討

percentage similarity はIPEM ToolboxやSetharesモデルとは方法論が全く異なる。これらの三和音の計算結果を比較してみると、percentage similarityはAugが期待通りのランキングに並び、異なる特徴がある。roughnessスペクトルの主なピーク周波数 (Bin No.) が調波系列となってお

り、そこに類似性があるのではないかと考えた。

4.5 類似性をみるための重み付け

Bin No.側のroughnessスペクトルは神経インパルスから抽出されたスペクトルで、その調波構造を見るためにpercentage similarityの近似曲線で重み付けを試みた。度合いの差も考慮した。重み付けの設定は入力音の根音(262(Hz))に基づくものを実験(1)、Pure ToneのMajでのroughnessスペクトルから抽出したものを実験(2)とした2種類を用意した。

4.6 類似性に関する実験結果

実験(1)の結果は、Organがうまく並ばなかった。一方、実験(2)の結果は、Pure Tone, Complex Tone, Piano, Organの各音色で三和音のランキングが並ぶ結果となった。

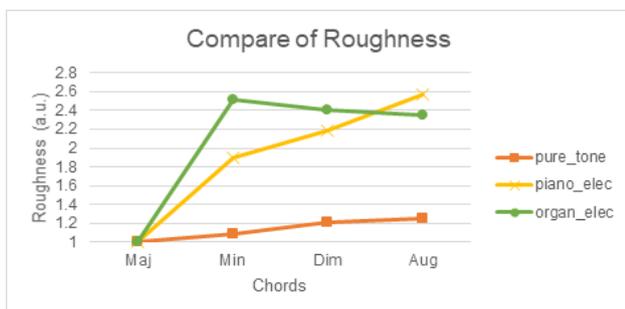


図3 実験(1)

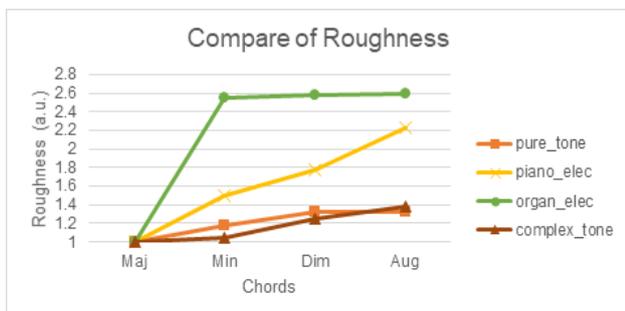


図4 実験(2)

4.7 Setharesモデルの適用の限界

C4, E4, G4の3音で構成されるMajに対し、E4を半音下げ、C4, E4 \flat , G4の3音で構成されるのがMinでMajより不協和度が上がる。さらにG4を半音下げC4, E4 \flat , G4 \flat の3音で構成されるのがDimで、G4からG4 \flat になることでMinに対して不協和度が少し上がる。ところが、MajのG4を半音上げたものがAugであるが、G4がG4 \sharp になることで不協和度が少し下がり、響きの良さはMaj < Augとなる。

4.8 percentage similarityの三和音への適用

MajとMinのpercentage similarityは、MajのC-E(長3度)、E-G(短3度)とMinのC-D \sharp /E \flat (短3度)、D \sharp /E \flat -G(長3度)と音程が入れ替わるだけの同じ組み合わせになるため

値が同じになる。

5. 結論

重み付けの実験(2)で、三和音(Pure Tone, Complex Tone, Piano, Organ)のランキングが並ぶ結果が得られ、roughnessスペクトルに類似性があることが確認された。さらに、roughnessスペクトル独自の調波構造を持つことが分かった。また、roughnessとpercentage similarityには相補的な関係があり、それぞれ単独で三和音のランキングは困難である。Johnson-Lairdの二重プロセス理論のような複合的なメカニズムを実験で試みることで、roughnessスペクトルとpercentage similarityとを相補的に組み合わせた協和度/不協和度を知覚する仕組みがあるという考えに至った。

文献

- [1] Helmholtz, H. *On the sensations of tone*. Margenau, H., Dover, New York (1954)
- [2] Sethares, W.A. *Tuning, Timber, Spectrum, Scale. Second Edition*, London: Springer-Verlag (2005)
- [3] 大串健吾著, 日本音響学会編: 音のピッチ知覚. 音響サイエンスシリーズ15, コロナ社 (2011)
- [4] 笠原潔, 徳丸吉彦: 音楽理論の基礎('07). 放送大学教育振興会 (2007)
- [5] 森 周司, 香田 徹 編, 日本音響学会編: 聴覚モデル. 音響サイエンスシリーズ3, コロナ社 (2011)
- [6] Fletcher, H. Auditory patterns. *Rev. Mod. Phys.* 12, 47-61(1940)
- [7] Moore, B.C.J. *An Introduction to the Psychology of Hearing Sixth Edition*. BRILL, LEIDEN • BOSTON (2013)
- [8] 入野 俊夫: 聴覚フィルタの測定と定式化について. 日本音響学会聴覚研究会資料, Vol.36 (2006)
- [9] Patterson, R.D., Nimmo-Smith, I. Off-frequency listening and auditory-filter asymmetry. *J. Acoust. Soc. Am.*, 67, pp.229-245(1980)
- [10] Glasberg, B.R., Moore, B.C.J. Derivation of auditory filter shapes from notched-noise data. *Hearing Research*, 47,103-138 (1990)
- [11] <http://dsam.org.uk/>
- [12] http://www.acousticscale.org/wiki/index.php/AIM2006_Documentation
- [13] Sethares, W.A. Local consonance and the relationship between timber and scale, *J. Acoust. Soc. Am.* Vol. 94, No. 3, Pt. 1 (1993)
- [14] 山本由紀子, 仁科エミ, 大西仁: 楽音を含む協和感と不協和感について. 『電子情報通信学会技術研究報告』, 113(299), 7-12 (2013)

- [15] Leman, M. Visualization and calculation of the roughness of acoustical musical signals using the synchronization index model (SIM). *Proceedings of the COST G-6 Conference on Digital Audio Effects (DAFX-00)*, Verona, Italy (2000)
- [16] Leman, M., Lesare, M., Tanghe, K. *Toolbox for perception-based music analysis Concepts, demos, and reference manual*, Version: 1.02 (beta) Date: 20140106
- [17] Van Immerseel, L.M., Martens, J.P. Pitch and voiced/unvoiced determination with an auditory model, *Journal of the Acoustical Society of America*; Vol. 91 Issue 6, p3511-3526 (1992)
- [18] Gill, K.Z., Purves, D. A Biological Rationale for Musical Scales. *PLoS ONE* 4(12): e8144. doi:10.1371/journal.pone.0008144 (2009)
- [19] Bidelman, G.M., Heinz, M.G. Auditory-nerve responses predict pitch attributes related to musical consonance-dissonance for normal and impaired hearing. *J. Acoust. Soc. Am.*, 130(3), 1488-1502 (2011)
- [20] Johnson-Laird, P.N., Kang, O.E., Leong, Y.C. On Musical Dissonance. *Music Perception, An Interdisciplinary Journal*, 30(1), pp.19-35 (2012)
- [21] McDermott, H.J., Lehr, J.A., Oxenham, J.A. Individual Differences Reveal the Basis of Consonance. *Current Biology* 20, 1035–1041 (2010)