

調和度に基づく和音性評価モデル -音楽ムードの色彩に基づく画像検索システムへの応用-

飛谷 謙介^{1,a)} 饗庭 絵里子^{1,2} 下斗米 貴之³ 谷 光彬⁴ 長田 典子¹ 藤澤 隆史⁵

概要：和音はメロディ・リズムと共に音楽を形作る重要な要素である。しかしながら、和音の物理的な特性と、そこから得られる知覚的な印象との関係については解明されていない部分が多い。先行研究として、和音性を不協和度・緊張度・モダリティで表す研究があるが、明るさという心理指標における検証が不十分であった。そこで、本研究においては、調和度という新しい和音性評価のための指標を定義し、和音の明るさの印象を推定する手法を提案する。また提案した手法を検証するために評価実験も行った。その結果、提案した調和度 (H) は心理実験の結果と強い相関が見られた。また、これらの結果について、音楽ムードの色彩表現に基づき画像を検索するシステムへの応用を行った。

Chord Character Evaluation Model Based on Harmoniousness: Application to Image Explorer System by Music Mood Color

1. はじめに

和音はメロディ・リズムと共に音楽を作る上で非常に重要な要素であり、音楽による様々な情動の喚起に欠かせない要因の一つである。和音によってもたらされる様々な印象は、和音を構成する音によって左右され、それぞれの和音ごとに固有の特徴が見られる。一般的には、長三和音は協和的な印象をもたらす、減三和音や増三和音は不協和な印象をもたらすという例がよく挙げられる。このような和音の心理的特徴と和音を構成する音の周波数比といった物理的特徴との関係について、多くの心理モデルが提案されてきたが ((cf.[1], [2], [3], [4]), 和音が明るく聞こえたり、暗く聞こえたりする要因となる物理量は未だ解明されてい

ない。そこで、本研究においては、和音の印象のひとつである明るさに関して、調和度に基づく新たな和音性評価モデルを提案する。また、これらの結果を応用し、音楽ムードを色彩によって表現し、その色彩によって画像を検索、表示するシステムの開発を行った。

2. 従来の和音性評価モデル

和音認知モデルの基本となる概念は2つの系統に大別できる。1つ目は和音の構成音が基音に対して(少なくとも近似的に)調和的な関係にあるという事実に基づくものである。提唱者の一人の Terhardt はこれを調和 (harmony) と呼び、協和と区別している [7]。2つ目は和音の構成音もしくは構成音の倍音のいずれかの組み合わせにおいてわずかなずれが生じている場合にうなりや粗さの感覚を生じ、不協和をもたらすという考え方である。これは Helmholtz に端を発するものとされており [8]、Plomp & Levelt が2音のインターバルから得られる感覚を不協和度として定量化した [1]。

Cook & 藤澤らによって提案された和音性評価モデル [9] もまた、この2つ目の概念に基づくものである。Cook らは、和音性を“和音から受け取る感性的性質の総体”と定義し、三和音を構成する二つの音程構造(根音 - 第三音、第三音 - 第五音)の周波数比から、不協和度 (D, 濁った一

¹ 関西学院大学 感性価値創造研究センター
Research Center for *Kansei* Value Creation, Kwansei
Gakuin University, Sanda, Hyogo, Japan

² 産業技術総合研究所
National Institute of Advanced Industrial Science and Tech-
nology

³ 玉川大学 脳科学研究所
Brain Science Institute, Tamagawa University

⁴ 関西学院大学
Kwansei Gakuin University, Sanda, Hyogo, Japan

⁵ 福井大学 子どものこころの発達研究センター
Research Center for Child Mental Development, University
of Fukui

a) tobitani@kwansei.ac.jp

澄んだ), 緊張度 (T, 緊張した-落ち着いた) 及び, モダリティ (M, 長調的な-短調的な) といった心理指標に対する理論曲線を算出している. 以下に式を示す.

$$D = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} v_{ij} \alpha_3 [\exp(-\alpha_1 x_{ij}^\beta) - \exp(-\alpha_2 x_{ij}^\beta)] \quad (1)$$

$$T = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} \sum_{k=0}^{n-1} v_{ijk} \exp\left[-\left(\frac{x_{ik} - x_{ij}}{\gamma}\right)^2\right] \quad (2)$$

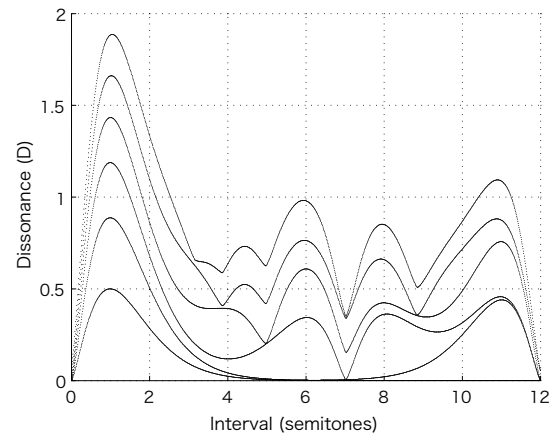
$$M = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} \sum_{k=0}^{n-1} -v_{ijk} \left[\frac{2(x_{ik} - x_{ij})}{\epsilon} \right] \exp\left\{-\left[\frac{(x_{ik} - x_{ij})^4}{4} \right]_{ijk}\right\} \quad (3)$$

ここで α , β , γ , δ , ϵ は定数, x は音程, n は計算に用いられる倍音の数, v は音量である. また, それぞれについて6倍音まで加算することにより, 以下のように拡張された理論曲線が得られる (図.1(a), (b), (c)). 図.1(b), (c) における根音 - 第三音間の音程は, 全て5半音である. これらのうち, 不協和度と緊張度を合わせたものから得られる不安定性という心理指標については, 行動実験によって得られたヒトの和音に対する印象評価結果に近いことが報告されている. しかしながら, 本研究で扱う心理指標である「明るさ」に最も近いモダリティについては, 増三和音, 減三和音および掛留和音 (suspended fourth) が同じモダリティを持つ和音であると算出されるなど, 問題が残っている. そこで, 本研究においては, 構成音の調和度に着目し, これを新たな和音性の指標となる物理量とし, 「明るさ」という心理量との関係性を導出する和音性評価モデルを提案し, その妥当性について検証を行う.

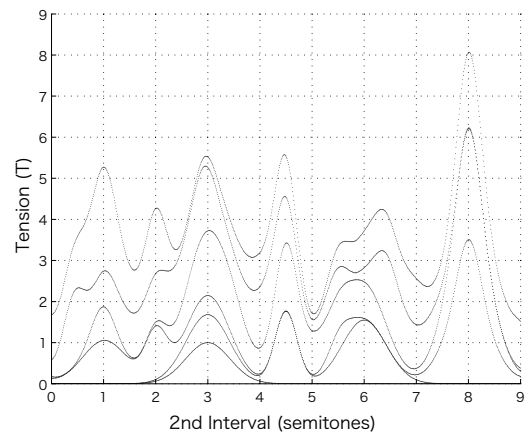
3. 調和度

3.1 和音の明るさ

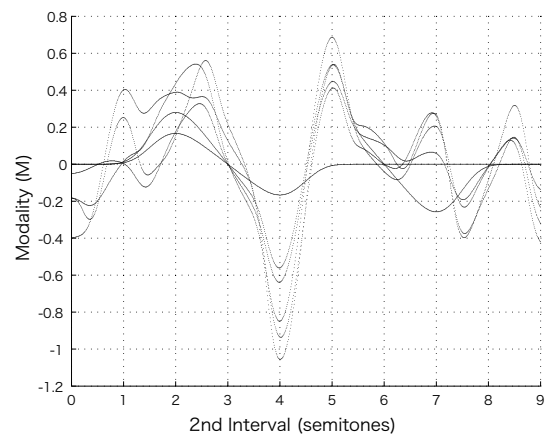
和音は, その種類によって固有の印象を喚起する. 特に和音の「明るさ」を指標とする評価は, 一般的によく言われるものであり, 例え3~4歳の幼児であっても, その違いを区別できることが明らかになっている [5]. 特に, 長三和音は「明るい」, 「嬉しい」, 短三和音は「暗い」, 「悲しい」といった印象や情動をもたらすことは, よく挙げられる例のひとつである. これらの情動は, 二音で構成される音程では喚起されないが, 三音で構成される和音になると喚起される. この現象に対する具体的な解答は得られていない. しかしながら, ピタゴラス音律をはじめとして, 音楽的な要素は周波数の比の単純さと関係があるとされてきた. 従って, 従来の和音性評価モデルにおいては, 二音間の音程を指標とすることが多かったが, その手法では和音を構成する三音の周波数比の関係を定量化することができ



(a) Dissonance Curve



(b) Tension Curve



(c) Modality Curve

図 1 Psychophysical Function of Chord Perception

ない. そこで, 本研究においては, 三つまたはそれ以上の数の構成音がいかに単純な整数比に近いかを算出する近似関数を用いることで, これらの課題を解決し, ヒトが和音の「明るさ」をモデル化する手法を提案する.

3.2 調和度の定義

本研究における調和度は, 構成音がいかに単純な比, またはそれに近い形となっているかということである. 従っ

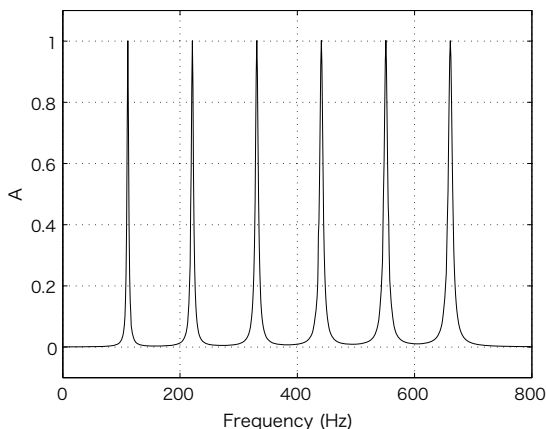


図 2 Approximate Function From Weighting of a Harmonic Overtone

て、調和度 $f(H)$ を算出するために、任意の音 f_0 の倍音付近に重み付けをした近似関数 $A(x, f_0)$ を用いた (図.2), 以下に、近似関数 $A(x, f_0)$ の式を示す.

$$A(x, f_0) = \sum_i \frac{1}{1 + \frac{N(x-if_0)^2}{f_0}} \quad (4)$$

N は、くし型を構成するそれぞれの山の裾の広がりを変化させる定数である.

この近似関数 $A(x, f_0)$ の f_0 を漸次変化させ、和音の周波数スペクトルである $f(x)$ との積分値を求め、これを類似度 $R(f_0)$ とする. この類似度 $R(f_0)$ が最大となる \tilde{f}_0 をもって、調和度 H とする. 以下に、それぞれの式を示す.

$$R(f_0) = \int f(x)A(x, f_0)dx \quad (5)$$

$$\tilde{f}_0 = \arg_{f_0} \max R(f_0) \quad (6)$$

$$H = R(\tilde{f}_0) \quad (7)$$

3.3 調和度の算出結果

主要な 3 和音 (図 3) におけるそれぞれの \tilde{f}_0 及び調和度の結果を表 1 に示す. ここでは $N=40$ とした. 調和度は maj. > sus4.=min. > aug. > dim. の順に低くなった. なお, Cook らのモダリティについての関数で算出すると maj. > sus4. = dim. = aug. > min. の順に低くなり, 全く異なる結果となった. そこで, モデルの妥当性を検証するために, これら五種類の三和音に関して, 「明るさ」についての評価実験を行った.

4. モデル検証実験：主要三和音の対比較

4.1 実験概要

本研究において, 提案した調和度 (H) の妥当性を検証

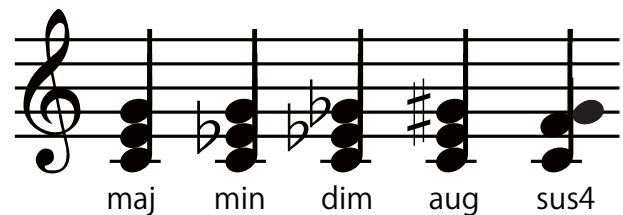


図 3 Principal Triads

表 1 Calculation Result of Harmoniousness

Chord	maj	min	dim	aug	sus4
Comp.1(Hz)	261.6	261.6	261.6	261.6	261.6
Comp.2(Hz)	329.6	311.1	311.1	329.6	349.2
Comp.3(Hz)	392.0	392.0	370.0	415.3	392.0
f_0	65.3	130.7	61.7	65.9	130.7
H	0.6992	0.6626	0.3842	0.4261	0.6627

するため, 図 3 で示した五種類の三和音を実験刺激とする一対比較実験を行った. 一対比較法においては, 二つの和音を連続して提示する必要があることから, 和音の進行によって喚起される印象, いわゆる順序効果をできるだけ抑えるために, 根音を C (261.6 Hz) で統一した. また, 第五音の高さが異なることによって受ける影響を避けるため, 根音のオクターブ上である C_1 (523.2 Hz) を追加した. 刺激対は, 全部で 20 通りであった. 和音刺激は MIDI を用いてピアノ音によって作成し, PC から再生したものをアンプスピーカーで提示した. 被験者は 22~25 歳の大学生 21 名である. 和音に対する印象のひとつである “明るさ” を評価五段階で評価した (“前の和音の方が明るい”・“前の和音の方がやや明るい”・“同じくらい”・“後の和音の方がやや明るい”・“後の和音の方が明るい”). 解析には, シェッフエの一対比較法の一つである芳賀の変法 [10] を用いた.

4.2 結果と考察

実験で得られた結果を図 4 に示す. 明るさの心理指標において, この結果から maj>sus4>min>aug>dim となることが分かった. これは 3.3 節で算出した調和度の順序と比較的よく一致している. 順位相関検定を行ったところ, 本心理実験による各和音の明るさの評価と, モダリティ関数によって算出された値の間には有意な相関が見られなかった ($\rho = 0.2, p > 1$, 順位相関検定). しかしながら, 本研究で提案した調和度に基づく評価モデルの間には, 強い相関がみられた ($\rho = 0.97, p < .01$, 順位相関検定). この結果は, 本提案モデルの妥当性を強く支持している.

5. 音楽の色彩表現システム

5.1 音楽の可視化

音楽ムードの可視化に関する研究は楽曲全体や 1 フレーズなどの大域的構造に注目したものと小節内などの局所

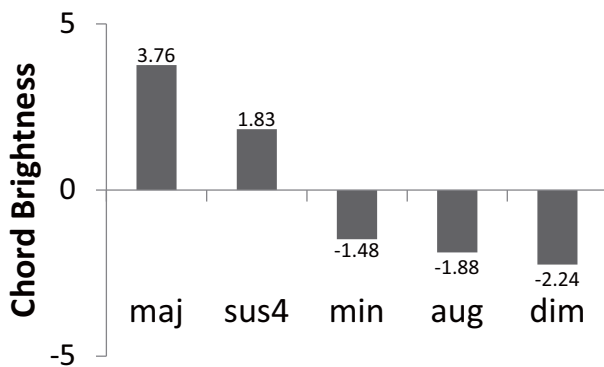


図 4 Evaluation Results of Chord Brightness

的構造に注目したものに大別できる。例えば“Shape of Songs”ではメロディのリフレイン構造から半円の重ね合わせによって楽曲構造を大域的に視覚化することができる [11]。後者では MIDI 音源などの入力に対して局所的なムードについて視覚化するが、その際に色相は楽曲の音階や調性へとカテゴリ的に割り当てられることが多い [12]。その際の問題として、以下の 2 点がよく挙げられる。1 点目は、調性概念の扱いである。調性概念は、音楽的要素(特に西洋音楽)の色合いが強く、多くの非音楽家は特に意識して聴取しているわけではないということである。2 点目は、調性と色相のマッピングについて、共感覚(色聴)など特殊なケースを除き [13]、未だ確証は得られていないという点である。そこで、本研究においては、音楽と色とのマッピングについて、調性概念ではなく和音性評価モデルの算出結果および心理実験によって得られた色と和音とのマッピングの結果を用いることとした。

5.2 色空間へのマッピング

和音性の 3 つの指標それぞれを任意の色空間に割り当てることにより、音楽ムードの色彩で表現を行う。色空間には、HSL を使用した。HSL 色空間は、色相 (H:Hue)、彩度 (S:Saturation)、明度 (L:Lightness) の 3 つの要素で構成され、人間の知覚により近いものとして、RGB 空間の代わりによく利用されるものである。先行研究における心理実験によって、不協和度 (D) が下がるほど、イメージされる色の彩度が上がることが示唆されている [14]。従って、不協和度と彩度を対応させることとした。同様に、調和度 (H) が高く明るい和音ほど、イメージされる色の明度が高くなったことから、調和度 (H) と明度を対応させることとした。次に、色相に関しては、彩度や明度と異なり、循環値であるため、まず色として分割した。分割の方法として、PCCS (Practical Color Co-ordinate System, 日本色彩研究所, 1964) を基準とし、赤、橙、黄、黄みの緑、青みどり、青、青紫、赤みの紫を採用した。心理実験によって、maj. や sus4. など、明るいと評価された和音に対し、黄みの緑、黄、橙などがマッピングされ、dim. や、

min., aug. において赤、青紫、青などの色がマッピングされていたことから、これらを採用した。

5.3 音楽の色彩表現による画像検索システム M-CUBE+

前節までに述べた和音性評価モデルを用いた音楽ムードの色彩表現手法を基に、その色彩を用いた画像検索システム (M-CUBE+) を開発した。本システムはプログラミング環境 processing によって実装し、その概念図を図 5 に示す。M-CUBE+ は和音推定モジュール、色彩マッチングモジュール、画像検索モジュールおよび画像出力モジュールの 4 つのモジュールから構成される。和音推定モジュールでは MIDI 楽器によって入力される MIDI 信号から和音を推定し、和音性評価モデルによって調和度や不協和度などを算出する。色彩マッチングモジュールでは前節に述べた方法により、算出した調和度や不協和度などに対応する複数の色を決定する。画像検索モジュールでは Flickr 画像データベースから求めた色の組み合わせによって画像検索を行う。画像出力モジュールでは、検索された画像群を様々な形で出力する。以上により、本システムは音楽とその音楽ムードの色彩の関係性を、検索された画像群を介すことによって、新たな形で知覚することを可能にする。

6. おわりに

本研究においては、和音性評価モデルにおける心理指標モダリティ (M) に代わる指標として、和音の明るさによって和音性を評価するための調和度 (H) を提案した。また、心理実験を通して提案モデルの妥当性を検証した。その結果、提案モデルが、ヒトの和音に対する明るさの評価を表現しうるものであることが示唆された。しかしながら、明るさの順序としては非常に妥当性が高いものの、maj. と sus4. および min. との間の差が比較的小さいことから、さらに改善の余地があると考えられる。また提案モデルをもとに音楽ムードの色彩による画像検索システムを開発した。本システムにより音楽とその音楽の持つ色彩感との関係性の新しい提示手法を開発した。また、今後の課題として、微分音を使用した和音など、より連続的な周波数によって構成された和音を評価可能なモデルを構築したいと考えている。これにより、ヒトが和音から明るさを感じる物理的要因をさらに詳細に検証することができれば、楽曲の印象から好みの楽曲や聞きたい楽曲を推薦するシステム、あるいは自動作曲システムなどにも汎用可能なモデルが構築できると考えられる。さらに、本システムは、単独の和音を評価対象としているが、和声進行によって進行後の和音の印象が左右される可能性もあることから、これらを考慮することで、さらにヒトの和音認知に近いモデルの構築を目指したい。

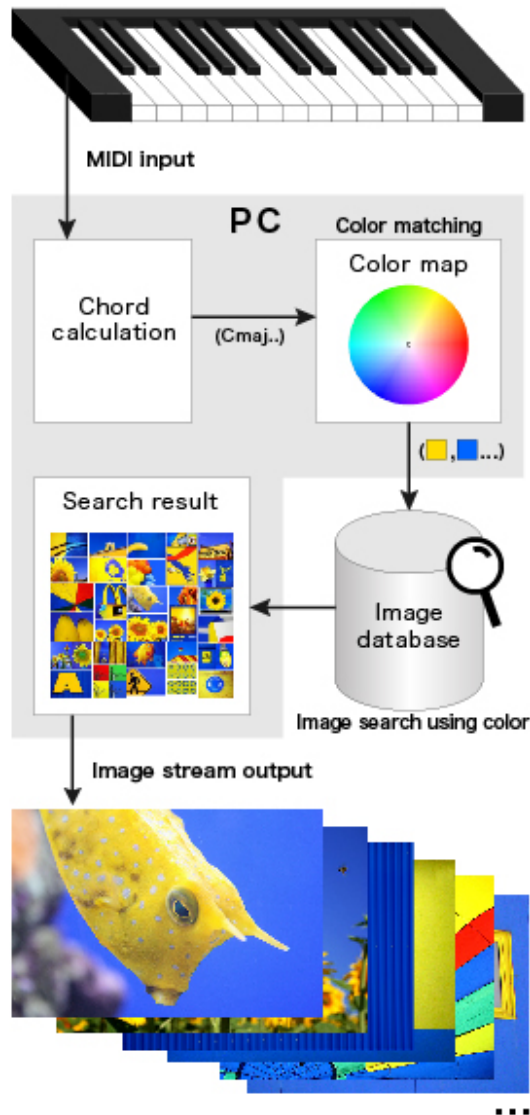


図 5 Diagram of M-CUBE+

参考文献

- [1] R. Plomp and W. J. Levelt, "Tonal consonance and critical bandwidth," *J Acoust Soc Am*, vol. 38, no. 4, pp. 548–60, 1965.
- [2] A. Kameoka and M. Kuriyagawa, "Consonance theory part i: consonance of dyads," *J Acoust Soc Am*, vol. 45, no. 6, pp. 1451–9, 1969.
- [3] —, "Consonance theory part ii: consonance of complex tones and its calculation method," *J Acoust Soc Am*, vol. 45, no. 6, pp. 1460–9, 1969.
- [4] R. Parncutt, *Harmony : a psychoacoustical approach*. Berlin ; London: Springer Verlag, 1989.
- [5] M. Kastner and R. Crowder, "Perception of the major/minor distinction:iv. emotional connotation in young children," *Music Perception*, vol. 8, no. 2, pp. 189–201, 1990.
- [6] T. Brown, S.-i. Tsujimura, A. Allen, J. Wynne, R. Bedford, G. Vickery, A. Vugler, and R. Lucas, "Melanopsin-based brightness discrimination in mice and humans," *Current biology*, vol. 22, no. 12, pp. 1134–1141, 2012.
- [7] E. Terhardt, "Pitch, consonance, and harmony," *J*

- Acoust Soc Am*, vol. 55, no. 5, pp. 1061–9, 1974.
- [8] H. v. Helmholtz, A. J. Ellis, and H. Margenau, *On the sensations of tone as a physiological basis for the theory of music*, 2nd ed. New York: Dover, 1954.
- [9] D. Cook Norman and T. Fujisawa X., "The psychophysics of hermony perception: Harmony is a three-tone phenomenon," *Empirical Musicology Review*, vol. 1, no. 2, pp. 106–126, 2006.
- [10] C. Sensory Evaluation, *Sensory Evaluation Handbook*. Tokyo: JUSE Press, Ltd., 1973.
- [11] M. Wattenberg, "Arc diagrams: Visualizing structure in strings," p. 110, 2002.
- [12] A. Mardirossian and E. Chew, "Visualizing music: Tonal progressions and distributions," in *In 8th International Conference on Music Information Retrieval*, Vienna, 2007, pp. 189–194.
- [13] N. Noriko, I. Daisuke, T. Manabu, H. W. Sanae, and I. Seiji, "Non-verbal mapping between sound and color : Mapping derived from colored hearing possessors and its applications," *The Transactions of the Institute of Electronics, Information and Communication Engineers. A*, vol. 86, no. 11, pp. 1219–1230, 2003.
- [14] T. Shimotomai, E. Aiba, T. Fujisawa X., N. Nagata, T. Omori, "Mapping Model from Chord to Color," *Proceedings of SCIS-ISIS2012*, pp.121-124, Kobe, Nov. 21, 2012.