

# 音楽における期待感の逸脱・実現の計算論的定式化の試み

大村英史 \*<sup>1</sup>    平田圭二 \*<sup>2</sup>    東条敏 \*<sup>3</sup>    柴山拓郎 \*<sup>4</sup>  
 Hidefumi Ohmura    Keiji Hirata    Satoshi Tojo    Takuro Shibayama

\*<sup>1</sup>東京理科大学 理工学部 情報科学科

Tokyo University of Science, Faculty of Science and Technology, Department of Information Sciences

\*<sup>2</sup>公立ほこだて未来大学 複雑系知能学科 知能システムコース

Future University Hakodate

\*<sup>3</sup>北陸先端科学技術大学院大学 情報科学研究系

Japan Advanced Institute of Science and Technology, School of Information Science

\*<sup>4</sup>東京電機大学 理工学部 システムデザイン学系

Tokyo Denki University, School of Science and Engineering, Division of Information System Design

Musical meanings based on realizations/deviations of an anticipation, on excitation/relaxation, or on stable/unstable. Some theorist proposed theories for these, such as a musical emotion model. However, we cannot implement these theories to a computer because these theories are based on ideational or psychological model. In this study, to implement realizations/deviations of musical anticipations to a computer, we attempt to model of realizations/deviations of musical anticipations using physical features of sound, information theory and probability.

## 1. はじめに

音楽聴取の楽しみは、期待感の逸脱・実現や、緊張・弛緩などによって成立していると考えられている [1]。このような音楽聴取のモデルや理論はいくつか提案されているが [2, 3]、概念的なモデルや理論にとどまっており、計算機上での実装は難しい。実際にはいくつかの研究では、音楽のジャンルに限定することにより分析可能なモデルを実現している（例えば、[4, 5]）。これらの方法の多くは、音楽に対してトップダウン的に分析を行う方法を用いている。つまり、現存する音楽や、それを生成するために発展してきた理論に基づいて分析を行う。音楽生成に伴う理論は、長い時間をかけて文化やジャンルなどに一般化されて純化してきた。そのため、これらの理論は局所的であると考えられることができる。例えば、和声理論をもちいると、古典音楽においては単一の分析結果を得ることができるが、近現代の音楽ではいくつかの分析結果が生じたり、異文化の音楽の分析はそもそもできない。このような理論を背景に、音楽理解のモデルを構築することは、効率のよい深い分析が可能とある。一方、人間が「さまざまな」音楽を楽しんで聴取する、といった根本的な人間の音楽聴取に関する原理まで到達できないのではないかと筆者らは考えている。この問題を乗り越えるために、音楽を構成する要素の基礎的な特性に着目して計算論的に再構築したモデルが必要であると考えられる。つまり、音楽に対してボトムアップ的に分析を行うモデルである。このようなモデルであれば、音楽の分析や生成だけでなく、音楽理論の分析や生成も可能となる。そしてこれらは、さまざまな文化やジャンルの音楽も包含できる可能性がある。本研究では、このようなモデルを検討する初期の段階として、音楽のパターンの根底にある物理的な特徴量、特に時間における相対的關係性と、それらの出現頻度でに着目する。本研究の最終的な目的

連絡先: 大村英史, 東京理科大学 理工学部 情報科学科, 〒 278-8510 千葉県野田市山崎 2641, hidefumi.ohmura@gmail.com

は、このような計算可能な基礎的な音楽聴取のモデルを作ることにある。

本稿の次のような構成となっている。まず、人間の音楽聴取の特性について考察をおこない、本研究で取り上げる物理的特徴量を紹介する。続いて、これら考察に基づいたモデルの提案を行い、このモデルに基づいた音楽生成システムを紹介する。最後に考察を行う。

## 2. 人間の音楽聴取

### 2.1 作曲と聴取の違い

楽譜は音楽を作り出すために用いられた道具である。そのため、音の高さや、発音タイミングなどだけではなく、拍子や調性などのパターンが記述されている。このような全体的な音楽的パターンの情報を、作曲家は事前に決定している。一方、聴取者は楽譜に記述されているような全体的な音楽の情報を即座に得ることはできない。ひとつひとつ音を聴取して、そこから音楽的なパターンを見つけ出す（見つけ出せない場合もあるかもしれない）。このパターンが期待となる。音楽の進行が、聴取者が得ることができたパターンに合致することを期待実現が生じる。そのパターンから逸脱していることを認識した場合は、期待逸脱が生じ、パターンの修正を行う。音楽の聴取のモデルを作るに当たり、楽譜に描かれるパターンを人間が理解しているのではなく、楽譜から作られる音の時系列的な変化を人間が理解していることに注意しなくてはならない。

作曲はパターン生成であるのに対して、音楽聴取はパターン認識とパターン修正である。そのため、音楽の聴取は作曲に比べて多くの人にとって容易である。もし、音楽の知識のある聴取者の場合は、既知の生成理論からパターンの先読みがなされることから、より複雑なパターン発見をすることも可能であろう。さらに、既知の音楽聴取の場合は、すでにパターンが既知であることから、ほとんどが期待実現であるが、未発見のパターンの発見によるパターン修正も生じることが予想される。

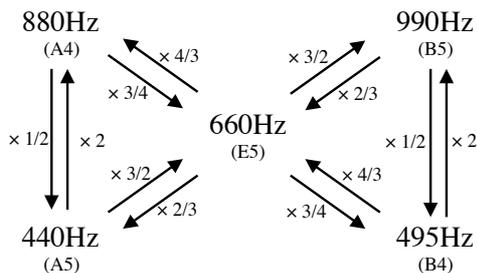


図 1: 音高同士の関係 (音程)

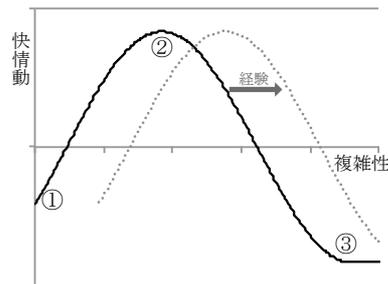


図 2: 最適複雑性モデル ( [] を変更)

本研究では、既存の理論は前提とせず、未知の音楽聴取に着目し、物理的な特性の知覚によってつくられるパターンに焦点を当てる。

## 2.2 パルスの関係 (リズム)

音楽聴取において、重要な基本パターンは反復である。反復がない構造は、音楽として感じられにくい (反復がなく音楽として感じられないこともまた、芸術の一環として取り上げられている)。リズムにおける最小の反復はパルスである。あるパルスとほかのパルスの周期の関係性で作られるパターンの認識がリズムの理解になる。

基本的に、これらの関係は素数の比によって作ることができる。最も簡単な関係は 1 倍である。1 倍の関係のパルスは同じ周期であるため、聴覚的にはパルスが 1 つの時と変わらない。2 倍の関係の 2 つのパルスを同時に聴取すると、2 拍子に聞こえる。続いて、3 倍の関係の 2 つのパルスを同時に聴取すると、3 拍子に聞こえる。5 倍であれば、5 拍子に聞こえる。しかし、5 拍子は、2+3 拍子や 3+2 拍子としてとらえられることを考えれば、人間にとっては比較的難しいリズムである。

次に、2 倍と 3 倍を合成することによって得られる、2/3 倍や 3/2 倍について考える。この関係によって作り出される 2 つのパルスを同時に聴取すると、どちらも 2 拍子と 3 拍子のポリリズム (複合リズム) に聞こえる。どちらか片方の周期を基準として考えると、特定の拍に裏拍が生じていると感ずることができる。3/4 倍や 4/3 倍の場合は、3 拍子 4 拍子のポリリズムが聞こえて、3 拍子と 4 拍子のどちらのリズムとしても感ずることができる。

実際の音楽は、より多くのパルスが混ざっている。この場合は、最も強いパルス (より頻度の高いパルス) を音楽の拍子として感ずることになる。また、それより小さな拍は裏拍や弱拍として感ずることになる。さまざまな音価がまじりあった単旋律の場合は、パルスの推定が難しいので、拍のパターンを感ずられないことも多い。

## 2.3 音高の関係 (音程)

音楽聴取では、音の高さ (音高) で作られるパターンも重要である。音楽理論では、音律論 [6] (音高の導出) や旋法論 [7] (音高の利用規則) としてまとめられている。音高は、空気振動の周波数である。実際の音は複数の周波数で構成されており、最も低い周波数 (基本周波数) がその音の音高として感ずられる。リズムと同様に、音高の関係性 (音程) は二つの音の周波数の関係で作られる。

ある音高の音と、2 倍もしくは 1/2 倍の音高の音は完全一度 (オクターブ) の音程をつくる。ある音高の音と、3 倍もしくは 1/3 倍の音高の音は完全十五度 (2 オクターブ) の音程を

つくる。これらは、最も協和している音程として感ずられるだろう。

次に 2 倍と 3 倍を合成して得られる、3/2 倍と 2/3 倍について考える。660Hz を基準にした場合 3/2 倍の音高は、660Hz と 990Hz<sup>\*1</sup> である。660Hz を基準にした場合 2/3 倍の音高は、660Hz と 440Hz<sup>\*2</sup> である。これは一般的に完全五度とよばれ、オクターブの次に協和している音程として感ずられるだろう。

4/3 倍や 3/4 倍について考える。660Hz を基準にした場合 4/3 倍の音高は、660Hz と 880Hz である<sup>\*3</sup>。660Hz を基準にした場合 3/4 倍の音高は、660Hz と 495Hz<sup>\*4</sup> である。これは一般的に完全四度とよばれ、完全五度の次に協和している音程として感ずられるだろう。ピタゴラス音律や三分損益法では、2 倍と 3 倍の関係性からあらゆる音を作り出している。

音高の関係の場合は、リズムと異なり 5 倍の関係も重要である。特に、3 つの音高の関係において、3:4:5 となる周波数比を構成すると、協和した音に感ずられる。この関係性を使った音律は、純正律である。

## 2.4 音楽における絶対量と相対量

リズムや音程のような関係を表す相対量は、ほぼすべての人が知覚できる。音楽を理解する上で、これら二つの特徴量は重要な要素であろう。

一方、パルスの周波数の絶対量の評価は困難である。つまり、速いのか遅いのかを感ずるためにはほかのパルスと比較する必要がある。音高においても同様である。その音が高いのか低いのかを判断するためにはほかの音の音高と比較する必要がある。これら絶対量は、一部の人間は比較することなしに知覚できる。しかし、多くの人は絶対的なパルスの早さや音高を識別することはできない。音楽の理解のためには、相対量は絶対量に比べて重要である。

## 2.5 音楽における複雑性

Meyer の期待逸脱理論にあるように、音の出現が期待を実現するか否かは、音楽における重要な要素である。期待を実現する音楽は複雑性が低く、期待を逸脱する音楽は複雑性が高いとして考えると、Berlyne の最適複雑性モデル [8] は参考となる (図 2)。このモデルによると、音楽の複雑性が低すぎたり、高すぎたりすると聴取者が不快に感ずる。音楽が適度な複雑性を保持するときに聴取者は快を感ずる。この複雑性を、人

\*1 音名で表すと E5 (ミ) と B5 (シ) (音律によって不正確)。

\*2 音名で表すと E5 (ミ) と A4 (ラ) (音律によって不正確)。

\*3 音名で表すと E5 (ミ) と A5 (ラ) (音律によって不正確)。

\*4 音名で表すと E5 (ミ) と B4 (シ) (音律によって不正確)。

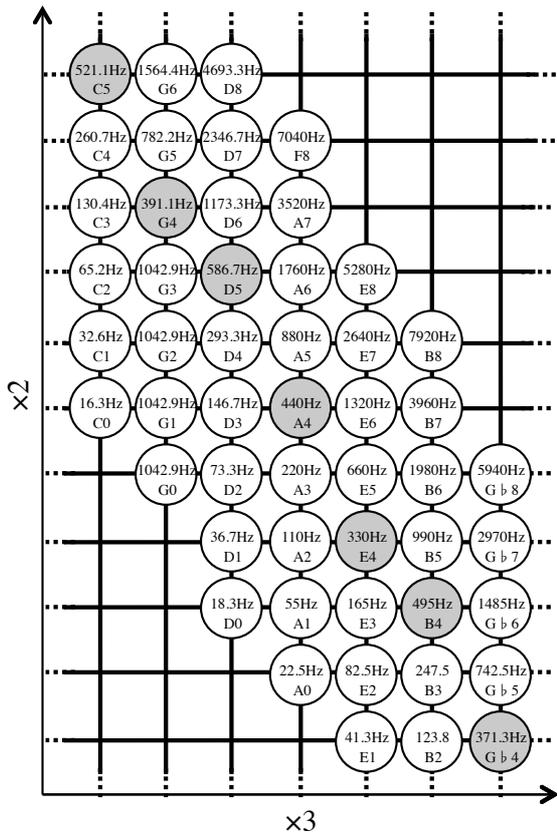


図 3: 2 倍と 3 倍で作る音高の格子空間

間を感じる不確実性としてとらえると, Shannon の情報理論 [9] におけるエントロピーによって定式化できる.

### 3. 音楽聴取のモデル

筆者らはこれまで, 音楽のリズムと音程のパターン生成のためにモデルを提案してきた [10]. このモデルでは, リズムパターンの生成は, パルスの周波数の 2 分割を任意の回数繰り返すことで作成している. つまり, 2 拍子系 (2 拍子や 4 拍子, 二分音符や四分音符など) のみを扱っており, 3 拍子系 (3 拍子や 9 拍子, 三連符など) を扱うことができない. さらに 2 と 3 の混合の拍子 (6 拍子など) も扱えない. 音程パターンに関しては, 五度圏に基づいている. 五度圏では, オクターブの関係性を同一視して  $3/2$  倍の比を使って 12 音の関係を規定している. 音程パターン生成のモデルでは, 五度圏に内在するピタゴラスコンマの誤差 (2 と 3 の累乗が一致しないことから生じる) は考慮に入れていない.

2.2 および 2.3 で考察したように, 音楽のリズムと音程のパターンは, 比による関係が関わっている. 素数倍の関係によってリズムと音程の基本的な要素は規定できる. しかし, 上記で考察してきたように旋律の作成という観点から, まずは, 2 倍と 3 倍の二つの関係を用いてモデルを作成する.

#### 3.1 2 倍と 3 倍の関係で構築する格子空間

まず, 音程における, 2 倍の関係と 3 倍の関係で成立するの二次元の格子状の空間を考える. 横軸を 3 倍, 縦軸を 2 倍とすると, 図 3 のように各周波数が配置される. この図は

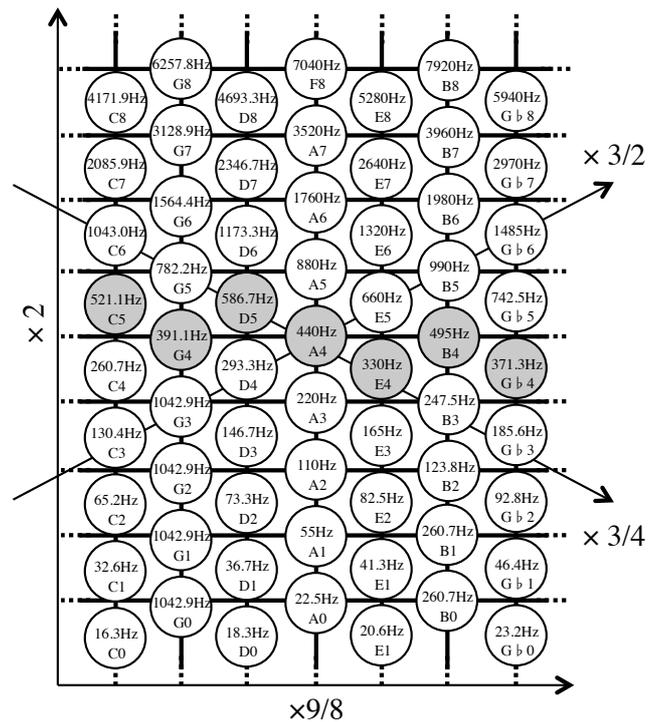


図 4: 図 3 を傾けた格子空間

440Hz を基準としている. 小数点以下第二位を四捨五入して記述してある. それぞれ, 周波数に対応した音高を併記してあるが, 基準 (ここでは 440Hz) から遠くなると誤差が大きくなり正確ではない. この空間は左斜め上方向 (右斜め下) に広がる空間であるため, 図 4 のように横軸を  $9/8$  倍になるように傾かせる. このようにすることにより横軸と縦軸に対してなだらかな空間となる.

続いて, リズムにおける, 2 倍の関係と 3 倍の関係で成立するの二次元の格子状の空間を考える. この格子における値の単位は bpm (beat per minute) とする. 36bpm を基準として横軸は 3 倍, 縦軸は 3 倍で各値を求めると図 5 のようになる. 36bpm を四分音符で表現したときの音符も表示している.

#### 3.2 格子空間内の分布

提案モデルでは, 各格子の頂点の出現確率を設定する. これを設定するために格子空間内に確率密度関数を設定する. ここでは, 2 次元のガウス関数を用いる.

$$f(x, y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \times \exp\left(-\frac{1}{2(1-\rho^2)}\left(\frac{(x-\mu_x)^2}{\sigma_x^2} - 2\rho\frac{(x-\mu_x)(y-\mu_y)}{\sigma_x\sigma_y} + \frac{(y-\mu_y)^2}{\sigma_y^2}\right)\right)$$

このとき,  $\mu_x$  は x 軸における平均値,  $\mu_y$  は y 軸における平均値,  $\sigma_x$  は x 軸における分散,  $\sigma_y$  は y 軸における分散,  $\rho$  は相関係数を表す.

音程に関する空間で, 混合分布を用いると, 長調, 短調のような旋法を表現することがわかっている [10]. そこで, ここでは二つの関数を用いた (図 6).

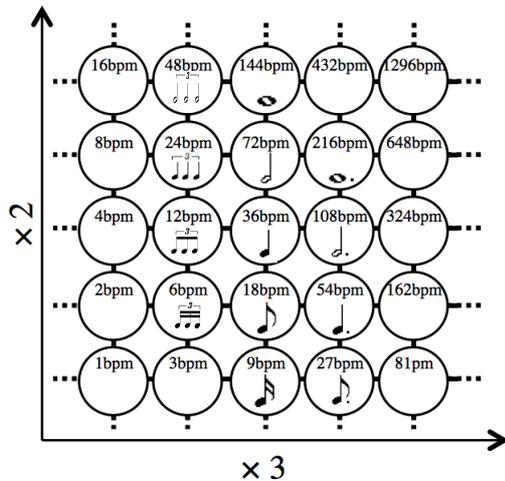


図 5: 2 倍と 3 倍で作るリズムの格子空間

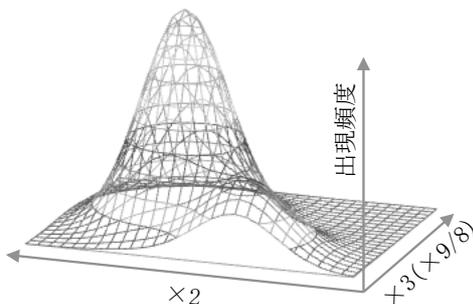


図 6: 格子空間内の出現確率

以上を、HTML および JavaScript で実装<sup>\*5</sup>した。

### 3.3 考察

作成したプログラムは、旋律を作成できる。関数の分散をコントロールして、旋律の複雑性をコントロールできる。出力を確認すると、快または不快な旋律が作成されているようには見え、旋律における複雑性の定量化は提案モデルで実現できそうである。しかし、実際の音楽が作成されたとは言えない。その理由は、音楽の階層構造理解の一番下の階層のモデルに過ぎないからだ。旋法や音律の構造と、旋律のリズム構造に限定される。私たちが聞いている音楽の分析や生成を行うためには、旋律同士の関係の複雑性も考慮に入れた階層的なモデルにする必要がある。

今回のモデルでは、5 倍以上の関係を省いているが、3 つ以上の旋律や和音には必要になる。このような旋律を考慮するためには、3 次元以上のモデルに拡張する必要がある。また、現在は、ピッチに関して、12 音で構成された音で音楽を作ることが当たり前になっている。そのため、素数倍の音高の音がほしい場合や、ある音高とほかの音高を動きのために間を保管する音がほしい場合に、その周波数に近い音高の音を近似

的に使っているに過ぎない。楽器の特性上、あらかじめ音高を用意しておく必要があるため、便宜上、12 音からそれに近い音を抽出しているように思える（ブルーノートスケールはこのような観点から微分音を使っているのかもしれない）。計算機上で音楽が作ることが可能になった現在ならば、12 音律にとられない音楽を作成できる。

今回は、音程とリズムに焦点を当てた。音楽の要素は、これだけではない。例えば、音色の倍音構造も複雑性でモデル化することにより、快不快の音色を作り出せるかもしれない。しかし、エンベロープなどのほかの音色の要素に、提案モデルを適用する方法は見当がつかない。

いずれにしても、提案モデルはまだ検討や拡張の必要性がある。今後は、実際の音楽の分析や、考察したような拡張に取り組んでいきたい。

## 4. おわりに

本研究では、音楽における期待感の逸脱・実現の計算論的定式化の試みを行った。音高の関係とパルスの関係から、音高およびリズムの複雑性を定量的に解釈する方法である。実装したシステムの出力を聴取すると、提案システムがうまく複雑性をコントロールできていることが確認できた。今後、モデルの拡充に取り組みたい。

## 参考文献

- [1] L. B. Meyer: *Emotion and meaning in music*, University of Chicago Press (1956).
- [2] E. Narmour: *The analysis and cognition of basic melodic structures*, University of Chicago Press, (1990).
- [3] D. Huron: *Sweet anticipation: Music and the psychology of expectation*, MIT Press (2006).
- [4] F. Lerdahl and R. Jackendoff: *A Generative Theory of Tonal Music*, MIT Press (1983).
- [5] A. Forte: *The Structure of Atonal Music*, New Haven and London: Yale University Press (1973). (森あかね訳: 無調音楽の構造: ピッチクラス・セットの基本的な概念とその考察, 音楽之友社 (2011))
- [6] 東川清一: 音律論: ソルミゼーションの探究, 春秋社 (2013).
- [7] 東川清一: 旋法論: 楽理の探究, 春秋社 (2010).
- [8] D. E. Berlyne: *Aesthetics and psychobiology*, Appleton Century Crofts (1971).
- [9] C. E. Shannon: *The Mathematical Theory of Communication*, The University of Illinois Press (1949). (植松友彦訳: 通信の数学的理論, ちくま学芸文庫 (2009)).
- [10] H. Ohmura, T. Shibayama, and T. Hamano: "Generative Music System with Quantitative Controllers Based on Expectation for Pitch and Rhythm Structure," *Proceeding of The Eighth International Conference on Knowledge and Systems Engineering (KSE2016)* (2016).

\*5 <https://sites.google.com/site/hidefumiohmura/home/program/jsai2017>