

研究報告

五度圏に基づいた確率的複雑化と偏向化による情動的旋法を用いた音楽生成システム

SOUND GENERATOR SYSTEM WITH EMOTIONAL MUSICAL MODE BASED ON PROBABILISTIC COMPLEXITY AND BIAS ON THE CIRCLE OF FIFTH

大村英史, 池田大樹, 吉池卓也, 金喜晴, 栗山健一

Hidefumi Ohmura, Hiroki Ikeda, Takuya Yoshiike, Yoshiharu Kim, Kenichi Kuriyama

独立行政法人 国立精神・神経医療研究センター 精神保健研究所 精神保健研究部

Department of Adult Mental Health, National Institute of Mental Health

National Center of Neurology and Psychiatry

概要

音楽は、期待の実現や裏切りによって情動を引き起こすと言われている。筆者は、このような期待の実現・裏切りを情報理論に基づいてリズムを生成するシステムを開発してきた。本発表では、この方法論ピッチ（五度圏）に適用させ、確率的分布の複雑性と偏向性を定量的にコントロールし旋法を生成する手法を提案する。そして、この手法を組み込んだ情動的サウンド生成システムを紹介する。

It is say that deviations and implications of expectation when listening to music arouses emotions. Following this perception, we developed the system which generated emotional rhythm using deviations and implications of expectation based on information theory. In this paper, I propose the method that is generating musical modes controlling probabilistic complexity and bias on the circle of fifth. I show the system generating emotional system with the method.

1. はじめに

音楽は、人間の情動に働く。普段の生活でも、多くの人間が情動コントロールに使っている [1]。音楽の人間への影響がわかれば、健常な人間の情動作用だけでなく、精神疾患の患者への応用も可能になる。さらに、抽象芸術全般や雰囲気といった事象なども、音楽と同様な性質を持ったものであると、筆者は考えており、応用可能である。これらを、定量化および定式化することができれば、上記にあげたような様々な分野の事象に活用が

可能である。

本研究では、このような抽象的な事象の定量化を目指し、まずは、音楽における物理的特性の一つである周波数の関係性に着目し、定量的なコントロールを試みた。

2. 音楽の特性

2.1. 二つの時間的特性

音は空気の振動である。一秒間に振動する回数を周波数としてあらわすと、この値が音の高さが決定する。この音の高さを音高 (pitch) という。一方、振動している (発音している) 期間を音価 (value) という。音価は一般的に、相対的な長さで表す。基準となる長さ (小節) に対する比で表すのが一般的である。両者とも、時間的な特徴量である。

音の特徴として、音色がある。音色は音の波形として考えることができる。任意の音を周波数解析すると、構成要素の正弦波を確認することができる。逆に考えると、音色は、異なる時間的な特徴量を保持した複数の正弦波から成り立っており、音高の時間的な特徴量の合成として見なせる。

また、音楽の構成要素の和音は、音高の異なる複数の音が同時に響くおとであり、音色と同様、音高の時間的な特徴量の合成として見なすことができる。

音楽の構成要素の律動 (リズム) は、基準となる時間的な長さ (たとえば小節) で作られるパルスであり、その中に音を配置する (たとえば裏拍) ことで複雑性が増す、音価の時間的な特徴量の構造である。

このように概観していくと、音楽の構造は、音高と音価の二つの特徴量に帰着することができ、時間的な関係

性によって成立している。

筆者はいままで音価に注目し、定量的にリズムを生成するシステムを開発した [2]。そこで、本研究では、もう一方の音高 (pitch) に注目し、定量的に音高を出力するモデルを提案し、システムに実装する。

2.2. 12 音階と 5 度圏

現在、世界中で最も多く用いられている音階は 12 音階である。この音階は、ピタゴラス音律や三分損益法と呼ばれる方法で得ることができ、基本的にはある音の周波数を 1.5 倍することで他の音を得ていく。12 音階は 1 オクターブ (2 倍の周波数) 内に 12 音を配置する。ある音の周波数を 1.5 倍を 12 回繰り返して得られる音が、2 倍して得られる音を 7 回繰り返して得られる音を等しいとして考えることによって作られる。しかし、 1.5^{12} と 2^7 は、わずかな誤差 (ピタゴラスコンマ) が生じるため、実際は近似的な 12 分割であり、これを修正するための方法はいくつか存在する。1 オクターブ (2 倍の周波数) 内に 12 音を配置する 12 音階はあくまでも近似であるが、1.5 倍 (5 度) と 2 倍 (オクターブ) でシンプルに構造化できるだけでなく、12 が多くの約数を持つため、多元的な関係を持った構造の構築が可能となる。

本研究では、12 音階を踏襲し、オクターブの関係は同一と見なし 5 度の関係と 4 度の関係 (5 度の逆) が最も近い関係であると見なす。つまり、5 度圏における両隣を最も近い関係だと見なす。

3. モデル化

3.1. エントロピー

Meyer によると、音楽における情動喚起は、音楽的期待からの逸脱によって生じる [3]。このような逸脱は、頻度の少ない事象として表現可能であり、情報理論でいうところのエントロピー [4] として表現可能である。

情報理論では、どの程度意味のある情報が伝達したのかを定量的に求めることができる。音楽は、時間変化と共に様々な構造を形成する音を聴取者が享受することにより、情動喚起を生じさせる。このような音の配置の伝達を、情報理論をもとにモデル化することにより情動喚起のための音楽を生成することが可能になると、私たちは考える。

情報理論は、事象 i が起きた際の情報量を次式で定義する。

$$I = -\log p_i$$

ここでは、イベント i が生じる確率を p_i とする。そして、 n 個の事象がそれぞれ、 $p_1, p_2, \dots, p_n$ で生じるとき情報量の期待値は次式で求められる。

$$H = -\sum_{i=0}^n p_i \log p_i$$

この値が大きくなると、多くの情報が得られることになる。つまり、不確実性が高まり複雑になると言い換えることができる。この値をエントロピーまたは平均情報量と呼ぶ。

12 音の各音に、出現確率を与えることで、上記のエントロピーを求めることが可能になる。

3.2. ガウス関数による各音の確率

12 音に各音の確率を与えるために、正規分布であるガウス関数を用いる。この理由は、2.2 で説明したように、音高には関係性があるからである。関係性が近いほど、音同士親和性が高く、離れると親和性が低くなると考える。

ガウス関数は以下の式で表される。

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

これは平均 μ 、分散 σ^2 の正規分布を表すグラフである。 x の値は各音に相当する。 μ と σ の値を変化させることで各音の出現確率を計算する。

3.3. 複雑性

ガウス関数の分散 σ^2 を小さくすると、曲線が扁平となる。すると、各音の確率がおおよそ等しくなり、一様分布に近づく。このとき、どの音が発音されるか予想することが困難で有り、複雑性が上昇している。一様分布になるとエントロピーは最大になる。一方、分散 σ^2 を大きくすると特定の音のみ発音されることになり、聴取者は次の音を予測することが容易になる。

システムでは、ガウス関数の面積が 95% に収まるような範囲で音が鳴るように設定されている。

3.4. 偏向性

発音範囲が決まったら、 μ の値により中央値をずらす。中央値は、最も頻度の出現頻度の高い音であり、主音として見なすことができる。右 (正) 方向にずらした場合、ガウス関数が 4 度の方向にずれる。左 (負) 方向にずらした場合は 5 度の方向にずれる。

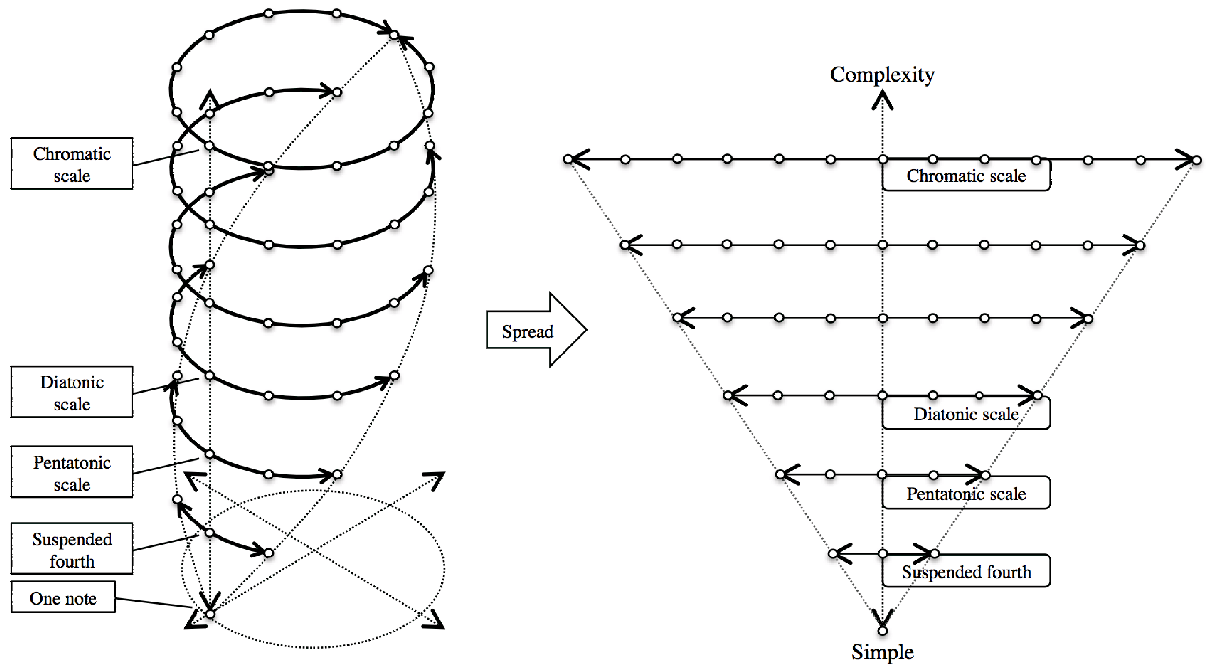


図 1. 複雑性の変化による選択される音

4. 実装

実装は、html 上に javascript を用いて行った¹。音の出力には Web Audio API を用いている。また、各パラメータをチューニングには midi 機器を利用できるように Web Midi を用いている。Google chrom で動作確認を行ってある。初期の Main ページで、音量や単位時間あたりの音の密度などが調整できる。音の密度設定に従って、ランダムで音出力される。左下の Play/Stop ボタンでシステムが起動/停止する。Setting ページでは、複雑性と偏向性をコントロールできる。また、手動でも各音の出現確率や、利用の可/不可がコントロールできる。

5. 考察

複雑性と偏向性をコントロールしてシステムを動かすと興味深い結果が得られる。複雑性および偏向性それぞれにおいて考察を行う。

5.1. 複雑性の考察

複雑性をコントロールしていくと図 1 のような変化をする複雑性を最小にすると、一音だけが選択される。この時はある音のオクターブだけが発音する。この値を

大きくすると 3 音が選択され、5 度と 4 度の音が発音される。この音は、コードの suspended 4th の響きと類似している。さらに大きくすると、5 音が選択され、ペンタトニックスケールのような響きとなる。さらに大きくすると、7 音が選択され、ダイアトニックスケールのような響きとなる。順に大きくしていくと 9 音、11 音となり、最終的に 12 音となり、クロマティックスケールと同等の響きが得られる。

複雑性を変化させることにより、選択される音が増え、複雑なサウンドが出力されていることがわかる。

5.2. 偏向性の考察

複雑性のコントロールにより、発音される音が決定されたら、中央値を変化させて、偏向性を変化させる。興味深いことに、5 度（負）の方向へ中央値を変化させた場合、明るい響きに変化していく。一方、4 度方向へ変化させた場合、暗い響きに変化していく。

例えば、7 音のときに教会旋法と比較すると、中央値を変化させない場合ドリア旋法のような響きであるが、5 度の方向へ中央値を移動させると、ミクソリディア、イオニア（長調）、リディア旋法のような響きに変化していく。各モードとも明るい響きである。一方、中央値を 4 度の方向へ変化させた場合、エオリア（短調）、フリギア、ロクリア旋法のような響きに変化していく。こちらは、暗い響きである。

5 音の場合も同様で、5 度方向へ変化させると、メ

¹ <https://sites.google.com/site/hidefumiohmura/home/program/entropy-mode>

ジャーペンタトニックや呂音階やヨナ抜き音階のような明るい響きになり、4度方向へ変化させると、マイナーペンタトニックや陽音階やのような暗い響きになる。

5.3. 情動的な旋法

以上をまとめると、複雑性と偏向性を変化させると、出現する音の組み合わせが変化して、情動的な響きの変化を得られることがわかる。Meyerの言うところの、期待の逸脱は複雑性（エントロピー）の変化に対応していると考えられる。一方で、長調・短調のような明るい・暗いと言った情動的变化は偏向性を変化させることで得られることがわかる。

提案したモデルでは、設定した二つのパラメータを変化させることで、旋法のような響きを得られたが、これらはすべて連続的に変化させることが可能だ。このような旋法は名称が付いており離散的だと考えられているが、2次元の連続的空間にマッピングできる可能性が示唆される。

6. おわりに

音の関係性により、出力される音を複雑性と偏向性の二つのパラメータでコントロールするモデルを提案した。このモデルを実装したシステムで出力されるサウンドを聴いてみると、実際の旋法と類似した響きを得られた。この結果から、旋法で定義されるような音の響きは、二つの特性によって作られる空間にマッピングすることができ、連続的である可能性が示唆された。このような定量的なコントロールは音楽が人間にどのような影響を与えているか、より厳密に調べることが可能となる。

今回は、5度の連続した関係のみ扱ったが、琉球音階や都節音階のような5度で連続していない音階も存在する。今後は、このような音階も考慮できるようなモデルも検討していきたい。

7. 参考文献

- [1] T. DeNora: *Aesthetic agency and musical practice, New directions in the sociology of music and emotion*, Oxford University Press, 2001.
- [2] H. Ohmura, T. Shibayama, T. Shibuya, T. Takahashi, K. Okanoya, and, K. Furukawa, "Modeling of Melodic Rhythm Based on Entropy toward Creating Expectation and Emotion," *Proceeding of Sound and Music Computing 2013 (SMC)*, pp.69-73, 2013.
- [3] L.B. Meyer, *Emotion and meaning in music*, Chicago: University of Chicago Press, 1956.
- [4] C. E. Shannon, *The Mathematical Theory of Commu-*

nication, The University of Illinois Press 1949. (植松 訳「通信の数学的理論」ちくま学芸文庫, 2009)

8. 著者プロフィール

大村英史 (Hidefumi Ohmura)

2002年東京農工大学工学部機械システム工学科卒業。2009年東京工業大学大学院博士課程修了。博士(工学)。同年JST, ERATO, 岡ノ谷情動情報プロジェクト研究員。および、理化学研究所脳科学総合センター生物言語チーム客員研究員、2014年国立精神神経医療センター精神保健研究所流動研究員、現在に至る。音楽によって生じる人間の情動の分析・応用に関する研究や、人工的雰囲気生成の研究に従事。知能情報ファジィ学会、人工知能学会、認知科学会、音響学会など会員。