

複雑性の変化による期待的情動のモデリングに 基づいた音楽的雰囲気生成

Modeling of Melodic Rhythm Based on Entropy toward Creating Emotion and Mood

大村英史^{*1,2} 柴山拓郎^{*3} 高橋達二^{*3} 澁谷智志^{*3} 岡ノ谷一夫^{*1,2,4} 古川聖^{*5}
Hidefumi Ohmura Takuro Shibayama Tatsuji Takahashi Satoshi Shibuya Kazuo Okanoya Kiyoshi Furukawa

^{*1}JST ERATO 岡ノ谷 情動情報プロジェクト
JST, ERATO, Okanoya
Emotional Information
Project, Japan

^{*2} 理化学研究所
Riken, Saitama,
Japan

^{*3} 東京電機大学
Tokyo Denki
University, Saitama,
Japan

^{*4} 東京大学
The University of
Tokyo, Tokyo,
Japan

^{*5} 東京藝術大学
Tokyo University of
the Arts, Tokyo,
Japan

音楽は期待感の連続とみなすことができる。聴き手は、音楽聴取時には次々に現れる音楽的事象がどのようなものであるのかということ推測している。その推測が正しいか正しくないかが、音楽聴取時における期待感を生み出す基礎となっている。しかし、期待感が達成される状態が連続すると、次第に聴き手は異なった状態を欲するようになる。したがって、快く期待を裏切ることが音楽情動の形成において重要であるといえる。これらのことから、期待感の裏切りには音楽構造の局所的な複雑さ(予測・外挿可能性)の増減が深く関わっていると考えることができる。しかし、この点に関する定量的な研究はまだ存在しない。本稿では、旋律の複雑さの定量化をエントロピーで表現する行うモデルの提案を行う。そして、このモデルが、複雑性を表現していることを検証した心理実験の報告を行う。

1. はじめに

音楽によって生じる雰囲気は音楽情動と言っても良いだろう。その理由は、音楽情動の研究は情動 (emotion) に限定したことでなく、感情 (affect), 気分 (mood), 情感 (feeling) なども含まれており、対象としている領域は広い [Juslin 2010]。最近では、日常生活において聴取者が気分をコントロールするためにどのように音楽が用いているか調べた研究も存在する [DeNora 2001]。雰囲気の研究の有名な図書である山本七平の「空気の研究」では雰囲気について様々な議論がなされている [山本 1983]。このなかで権力という用語の重要性が読み取れる。構成されるメンバーによって作り上げられる雰囲気があるが、権力を持った者によって水が差されその雰囲気が一気に崩壊し新たな雰囲気が生じるということだ。音楽が「音による芸術」という広辞苑による意味を考慮して、これを当てはめると、前者は音楽の要素、つまりメロディやリズムや音色などの音による音楽的構造によって多くの聴取者が感じる情動を指す。これはオーディオプレイヤーなどに代表される演奏行為から完全に分離した音という要素によって生じる情動が良い例である。一方、後者は音楽がもつ背景から得られる情報によって作り上げられる雰囲気を指す。例えば、ビートルズのコンサートで音楽を聴く前に失神してしまうような雰囲気などが良い例である。本研究では、権力によって作り上げられる雰囲気(ここでは特に商用音楽に代表されるような音楽以外のメディアを通して価値付けされる雰囲気)ではなく、音楽のみによって成立する雰囲気(音楽的構造であるリズム、ピッチ、音色などによって作り上げられる情動)を対象にする。

音楽聴取による情動喚起は、音楽研究における重要なトピック

連絡先: 大村英史, JST ERATO 岡ノ谷情動情報プロジェクト,
〒351-0198 埼玉県和光市広沢 2-1 理化学研究所 研究
本館 411, tel: 048-462-1111(内線 3452), E-mail: ohmura
@brain.riken.jp

クである。Meyer は、「複数の傾向 (tendency) の対立によってブロックないし抑制によって感情が喚起される」とする Dewey の情動の葛藤理論 [Dewey 1894] を援用して、音楽的期待からの逸脱が情動を喚起すると指摘している [Meyer 1956]。Narmour は期待 (implication) に対する実現/裏切り (realization) の関係を IR 理論としてまとめている。このような期待に関する研究は Huron も ITPRA 理論にまとめており [Huron 2007]、期待の実現と裏切りが音楽における情動に関わっていることは間違いないと考えられる。これらの、実際にまとめ上げられた理論では、期待の生成や実現/裏切りは、直感的な感覚に基づいたルールを基盤として定義されている。そのため、これらの理論を直接計算機に実装することは困難である。私たちは情動的音楽の自動生成をゴールに設定し、期待および実現/裏切りを生成するための計算モデルを提案する。音楽的期待からの逸脱は、予め期待された構造の規則性(パターン)の減少あるいは消失する状況から引き起こされる。この状況は、不確実性が増加していることから、情報理論的にエントロピーが増加し複雑化していると思える。このような不確実性は、複雑性と快楽の関係について述べている Berlyne の最適複雑モデル [Berlyne 1971] の「複雑性」と共通する点がある。これは、不確実性と情動には関係があることを示唆している。このように、エントロピーを用いることによって、音の配置から不確実性を計算することが可能となる。(個々わかりにくい) また、逆にエントロピーに設定により、目的の不確実性をもった音の配置が可能となる。(ここまで) 本研究では、メロディにおけるエントロピーの操作によって情動喚起を促すことを指針としている点の特徴である。一般的にメロディは音高(音の高さ)と音価(音の長さ)の集合として記述することができる。また、音色もメロディの特徴を表す要素の一つである。私たちはこれまでに音高に限定したメロディ生成のシステム開発を行ってきた。本稿では、次のステップである音価について着目し、メロディのリズムを生成するための基準となる、リズムの不確実性を扱うモデルの提案を行う。そして、このモデル

に基づいたメロディ生成システムを作り、そのメロディの評価実験を行うことで、モデルの妥当性を検証する。

2. 音楽的構造と複雑性

2.1 音楽的期待を作る音楽構造

Meyer は、聴取者が音楽によって作られる期待が裏切られたときに情動が生じると述べている [Meyer 1984]. Narmour は、Meyer のこの考え方に基づき、音楽聴取中に作られる期待を「暗意 (implication)」として I-R 理論 (implication-realization theory) を提案している [Narmour 1990]. 連続する 2 音は、特定の関係を導き、次の音の予測を作ることができる。実際に提示された音がその予測をどのように実現したかによって聴取者は感覚を持つ。この感覚が情動を引き起こす。つまり、連続する 2 音が「暗意 (implication)」を作りだし、後続する音が「暗意」に対する「実現 (realization)」をもたらす。私たちはこの考え方を基に、音の高さの関係性を用いて人間の認知的なバイアスである対称性バイアスから遷移確率を作り出しメロディ生成を試みた [大村 2012]. その結果、音楽的なまとまりや音楽的な意外性を聴取者に感じさせるメロディ生成を実現した。このようなメロディを分析した結果、まとまりや意外性の感覚は複雑性を表すエントロピーと関係していることが明らかになった。もちろん、この関係性は音の高さだけでなく音の長さで作り出すリズムにもあると考えられる。

2.2 複雑性を表すエントロピー

期待の生成や実現裏切りにおけるまとめられた理論は、直感的な感覚に基づいたルールによって作られている。そのため、計算機におけるこれらの実装は適していない。私たちは情動を基盤とする音楽の自動生成システムの構築をゴールとして設定し、期待および実現/裏切りを生成するための計算モデルを提案する。そのために、期待および実現/裏切りを実現するために情報理論 [Shannon 1948] に着目した。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教授することにより、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。

$$I = -\log p_i \quad (1)$$

ここでは、 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。

$$H = -\sum_{i=1}^n p_i \log p_i \quad (2)$$

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。

3. 提案モデル

3.1 メロディにおけるリズムの作り方

本研究ではメロディに焦点を当てて、メロディ生成を目指している。メロディの性質は、要素である音価(音の長さ)、音高(音の高さ)、音色 (timbre) などいくつかの要素で決まる。私たちは

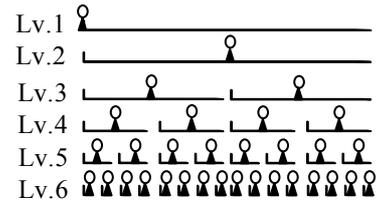


図 1 分割方法

表 1 $\sigma = 10$ の時の各レベルの重みと確率

	Lv. 1	Lv. 2	Lv. 3	Lv. 4	Lv. 5	Lv. 6
x	0	12	24	36	48	60
w	9978	4855	560	16	1	1
p	0.647	0.315	0.036	0.001	0.000	0.000

音高の関係性によるメロディ生成を試みた。その結果、音高の遷移のエントロピーによって、人間の複雑性の知覚を作ることができることが示唆された。本稿では、次のステップとして音価のコントロールによる複雑性リズムのモデルの提案を行う。最小のリズムのパターンは、任意の時間領域の反復によって作られるパルスである。パルスから、音価を変化させることによってつく。一般的にリズムは任意の時間領域の反復によって作られる。これを拍節という。一つの拍節の中で、どのような位置に音を配置するかを決定する。これらの位置は拍節を分割することで得る。分割は素数での分割を繰り返し行うことで、様々な位置を得ることが可能である。音楽では、2 分割、3 分割を用いることが多い。5 分割以上は複雑になるためほとんど用いられることはない。例えば 5 拍子の曲を 2+3 として捉えたり、7 連符の演奏を 4+3 として捉えたりする。これは、5 以上の分割は人間にとっての知覚は困難であるからだろう。また、分割が混合している場合も人間は性格に知覚ができないと言われている(Essens 1986)。このことから、リズムパターン生成の基本は 2 または 3 分割を考えれば良いことが示唆されている。今回は、単純化のため、2 分割のみ扱う。

図 1 に分割方法の概要を示す。Lv. 1 は分割なしの位置を示す。Lv. 2 は 1 回分割を行った位置を示す。Lv. 3 は 2 回、Lv. 4 は 3 回と示し、Lv. 6 まで 5 回までの分割を示す。例えば、拍が 4/4 の場合、Lv. 1 は全音符を組み込む場合の位置であり、Lv. 6 は 32 分音符を組み込む場合の位置である。Lv. 7 以上も考えることが可能であるが、64 分音符以上は装飾音との区別がつかなくなることが予想され、リズム的要素として意味をほとんど持たないと考え、Lv. 6 までを考慮することとする。これにより、拍節内に 32 個の位置が存在し、ここから選択する事でメロディのリズムを生成する。

3.2 エントロピーの作り方

エントロピーは 32 個の位置をレベル分けすることで、入りやすさを確率で与えることで得る。Lv. 1 から Lv. 6 に、それぞれ重みをあたえる。この重みは、より低いレベルに大きな値を与える。その理由は、高いレベルは一つ下のレベルの音を分割することによって得られるため、レベルが高くなるにつれて聴取者が認識しづらいつけられるからだ。

重みの傾斜は以下の式で表されるガウス関数で与える。

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \quad (3)$$

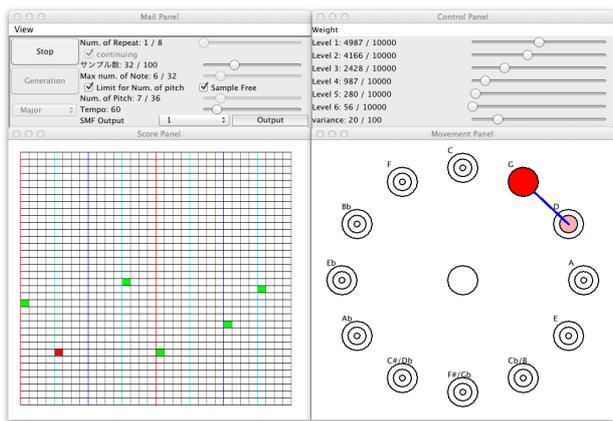


図2 提案モデルが実装されたアプリケーション

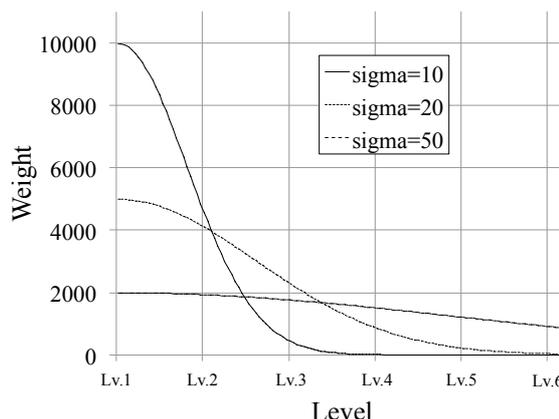


図3 実験で用いた3つのガウス曲線

表2 実験で用いた各値

σ	p						H
	Lv.1	Lv.2	Lv.3	Lv.4	Lv.5	Lv.6	
10	0.647	0.315	0.036	0.001	0.000	0.000	1.115
20	0.389	0.323	0.188	0.076	0.022	0.004	1.947
50	0.210	0.204	0.188	0.163	0.133	0.102	2.544

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 の値はレベルに相当する。 μ の値は 0 に設定し、 σ の値を変化させることで得られ値を各レベル重みとする。今回は重みを見やすくするために $w = f(x) \times 2.5 \times 10^5$ (if $f(x) < 1$ then $f(x) = 1$) とした。例えば、 $\sigma = 10$ のとき、表 1 のように各レベルの x を設定すると、重み w , 確率 p が求められる。この確率を式(2)に代入することにより、エントロピー $H = 1.115$ が求められる。ここから生成されるリズムを持ったメロディの例は 4.3 実験メロディで記述する。

3.3 モデルの実装

実装は Java で行った。図 2 のような 4 つのパネルからなるインタフェースを持つ。左上の Main Panel では同一拍節の反復回数、拍節の自動連続生成による再生、音を探すためのサンプル数、拍節内の上限音符数、上限音符数に達したときに探索を中止する打ち切り探索チェックボックス、上限音符数に満たない場合は指定サンプル数を無視して上限音符数まで達するまで探索を指示するサンプルフリー探索チェックボックス、規定の音高数からランダムで音高が選択されるが、その数を設定する音高数、テンポ、音高を取得する音階の設定、そして SMF の出力からなる。右側の Control Panel では各レベルの重みを個別に設定できる。また、一番下の分散を設定するとその値にあった各レベルの重みが自動的に設定される。

Score Panel では縦軸が音高、横軸が時間の 32×32 のシーケンサーとして、生成されたメロディを表示する。Movement Panel では再生時に音高がわかるように 5 度圏で表された音高の位置を赤で表示する。円の大きさは高さで有り、小さい円が高く大きい円が低いことを表している。

4. 実験

4.1 実験概要

本実験の目的は、エントロピーを変化させて生成されたメロディが聴取者に違いをもって認識されていることを確認することで



図4 メロディの例

ある。実験参加者は 25 人であったが、全て回答が同じであった者や、一般的な楽曲との比較を行った者など実験意図を伝達しきれなかった 5 人を分析から抜いた。そのため、20 人(男性 11 人、女性 9 人)、平均年齢は 31.57 歳(標準偏差:9.11)の結果に対して分析を行った。

4.2 実験システム

実験は、PHP および javascript で書かれた実験用 web ページで行った。各実験参加者は任意の計算機の web ブラウザから実験用ページにアクセスし、ページ内にあるプレイヤーでメロディを再生・聴取し提示されている項目に対する評価を行った。実験は、イヤホン・ヘッドホンを用いるか静かな環境において、適切だと感じる音量で行うように指示された。

4.3 実験メロディ

3 つの条件 ($\sigma = 10, 20, 50$) でメロディを生成する。各条件のガウス曲線は図 3 に示すとおりである。これらの確率、およびエントロピーは表 2 に示す。生成されるメロディをそれぞれ、メロディ 1、メロディ 2、メロディ 3 と呼ぶ。メロディの長さは 8 拍節でおよそ一分である。今回は全てのメロディにおいて 1 拍節内に 6 つの音を配置させた。実験参加者はそれぞれの条件で作られたメロディを聴取するが、各実験参加者に対して新しく聴取する 3 つのメロディを作る。そのため、実験参加者によって聴取するメロディは異なる。生成されるメロディの例を図 4 に示す。実験では、聴取する各条件のメロディの順による影響を避けるためランダムにしてある。

4.4 評価項目

メロディの感覚的知覚の評価として以下の項目を質問した。括弧の中は質問項目の省略形で、以下ではそれらを用いる。実験参加者は各項目を1から5の5段階で評価した。1はまったく感じないことを表し、5は強く感じる事を表している。

- リズムを感じる(リズム)
- 音楽的にまとまりがある(まとまり)
- 複雑だと感じた(複雑)
- 音楽的面白さがある(面白さ)
- ポジティブな情動が引き起こされる(ポジティブ)
- ネガティブな情動が引き起こされる(ネガティブ)
- このメロディが好きだ(好き)

評価項目の順番による影響を避けるために、実験参加者ごとにランダムで順番を生成した。

4.5 実験手続き

実験は以下の手続きにより実施した。

- 1 実験参加者は実験 web ページに任意の計算機からアクセスする。
- 2 実験説明文をよみ終わったら、年齢、性別などのアンケートに答える。
- 3 メロディを聴取し評価を行う。
- 4 全ての回答が終わったら、データを送信して終了する。

4.6 実験結果

各質問項目における評価値の分散分析を行い、平均値の差の検定を行った。平均値(標準偏差)と分析結果を表3に示す。

「リズム」、「まとまり」、「複雑」において有意な差があったので、これらを多重比較(Bonferroni)で分析した。「リズム」においては、メロディ1よりメロディ3($p=0.003$)、およびメロディ2よりメロディ3($p=0.180$)の方が高い評価をした。「まとまり」においては、メロディ1よりメロディ3($p=0.037$)を低い評価をした。「複雑」においては、有意性は出なかったが、メロディ1よりメロディ3の方が高く評価する傾向にあった($p=0.055$)。

「面白さ」「ポジティブ」「ネガティブ」「好き」においては、有意な差が無かった。

5. 考察

エントロピーの操作によって「リズム」、「まとまり」、「複雑」の評価を変えることができた。これにより、私たちの提案するモデルの妥当性が検証された。メロディ1とメロディ3は3つの評価において全て差があった。しかし、メロディ1とメロディ2の差と、メロディ2とメロディ3は差が出にくい傾向にあった。この理由はエントロピーの差が小さいときには、聴取者に差を感じさせることができないことが予想される。ある程度エントロピーに差があるときに、聴取者は差を感じ取ることができる。この差を感じることができる閾値がどのあたりにあるのか、さらなる検証が必要だ。特に「複雑」の評価に関しては差が出にくかった。これは、メロディのリズムをつくる音価以外の要素である音高の影響があったことが予想される。音価と音高の影響は独立ではなく、相互作用を与えている可能性が高い。

一方、「面白さ」、「ポジティブ」、「ネガティブ」、「好き」の評価はエントロピーの操作によって変えることができなかった。情動や嗜好といった主観的な評価は、エントロピーが一定の場合には生じないことがわかる。今後、メロディ進行中のエントロピーの操作によって、期待を作り出し、その期待の実現/裏切りを作っていく予定である。

表3 各評価の平均値と分散分析 * $p<0.05$, ** $p<0.01$

メロディ		1 ($H=1.115$)	2 ($H=1.947$)	3 ($H=2.544$)
リズム	Mean (SD)	3.4 (0.995)	3.15 (1.137)	2.25 (1.020)
	ANOVA	$F(2, 57) = 6.609, p = 0.003^{**}$		
まとまり	Mean (SD)	2.85 (0.988)	2.55 (0.999)	2.05 (0.945)
	ANOVA	$F(2, 57) = 3.420, p = 0.040^*$		
複雑	Mean (SD)	2.85 (1.127)	2.65 (1.182)	3.55 (1.050)
	ANOVA	$F(2, 57) = 3.262, p = 0.046^*$		
面白さ	Mean (SD)	2.8 (1.005)	2.75 (0.786)	2.85 (1.226)
	ANOVA	$F(2, 57) = 0.479, p = 0.953$		
ポジティブ	Mean (SD)	2.45 (1.146)	2.45 (0.887)	2.55 (1.356)
	ANOVA	$F(2, 57) = 0.508, p = 0.951$		
ネガティブ	Mean (SD)	2.2 (1.005)	2.1 (1.071)	2.45 (1.099)
	ANOVA	$F(2, 57) = 0.579, p = 0.564$		
好き	Mean (SD)	2.2 (0.952)	2.6 (1.273)	2.4 (1.429)
	ANOVA	$F(2, 57) = 0.525, p = 0.594$		

6. まとめ

私たちは、音楽における期待の実現/裏切りを利用した、情動的音楽の自動生成システムの開発に取り組んでいる。本稿では、期待の実現/裏切りを作り出すためのメロディの定量化としてエントロピーに着目し、リズムにおける複雑性を作るためのモデルを提示し、そのモデルの妥当性を実験で確かめた。今後は、エントロピー操作で実現した複雑性を用いた期待の実現/裏切りを生成する手法を確立していきたい。

謝辞

本研究は東京電機大学総合研究所研究助成 Q11J-08 および、JSPS 科研費 24603007 の助成を受けた。

参考文献

- [Juslin 2010] P. N. Juslin and J. A. Sloboda: Handbook of music and emotion, theory, research, application, Oxford University Press, 2010.
- [DeNora 2001] T. DeNora: Aesthetic agency and musical practice: New directions in the sociology of music and emotion, Oxford University Press, 2001.
- [山本 1983] 山本七平: 空気の研究, 文藝春秋, 1983
- [Dewey 1894] J. Dewey: The theory of emotion: I: Emotional attitude. Psychological Review, 1 (6), pp. 553-569, 1894.
- [Meyer 1956] L.B. Meyer: Emotion and meaning in music. Chicago: University of Chicago Press, 1956.
- [Huron 2007] D. Huron: Sweet anticipation: Music and the psychology of expectation. Cambridge, MA: MIT Press, 2006.
- [Berlyne 1971] D. E. Berlyne: Aesthetics and psychobiology. New York: Appleton Century Crofts, 1971.
- [Narmour 1990] E. Narmour: The analysis and cognition of basic melodic structures. Chicago: University of Chicago Press, 1990.
- [大村 2012] 大村英史, 柴山拓郎, 高橋達二, 澁谷智志, 岡ノ谷一夫, 古川聖: 人間の因果推論による認知バイアスに基づいたメロディ生成システム, 日本知能情報ファジィ学会誌「知能と情報」, 24 (5), pp954-966, 2012.
- [Shannon 1949] C. E. Shannon: The Mathematical Theory of Communication. 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)
- [Essens 1986] P. Essens: Hierarchical organization of temporal patterns. Perception & Psychophysics, 40, pp. 69-73, 1986.