

地図標高データを用いたメロディ生成法

梅村 祥之^{a)} 富士 直斗²

概要: 自然界に存在するパターンから曲を生成する技術として、大量かつ容易に入手できる地形の標高データを使って主旋律を生成する技術を開発した。初心者の音楽制作支援用にポップス曲を想定した曲生成アルゴリズムである。標高パターンに一般の音楽の持つ特徴を加えるにあたって、音楽理論に基づくルールベースの処理方法を採用した。生成曲の評価として、音楽性を隣接音符間の音程の頻度分布などの統計量により確認した。人間の作曲による既存曲5曲と生成曲1曲を使って、好みの程度を主観評価する実験を行い、有意差検定することにより、生成曲が既存曲のうちの2曲よりも有意に好まれるという結果を得た。

キーワード: 自動作曲, ルールベース, 作曲支援, 旋律, 主観評価

A method to generate melodies using altitude data of mountains

YOSHIYUKI UMEMURA^{a)} NAOTO FUJI²

Abstract: We have developed a method that generates melodic lines using curves of altitude data provided in large quantities representing the shape of mountains. The algorithm is made on the assumption that it is used as an assisting system for beginners to create J-pop style tunes. The processing that include properties of music into altitude sequences is a rule-based processing. We verified musicality of generated tunes using statistical features such as a distribution of intervals between neighboring notes. We experimented subjective evaluations of preference for a generated tune and five human-made tunes. As the result, the generated tune was preferred to two human-made tunes significantly.

Keywords: automatic composition, rule-based, assistance of composition, melody, subjective evaluation

1. はじめに

自動作曲技術の研究の中で、自然界に存在する信号、もしくは自然界の形状を模擬した生成モデルによる信号を基にして曲を生成する方法や、自然物の中でも声のように人の発する信号を基に曲を生成する方法などが研究されている。1つ目のタイプは、自然界のパターンに音楽が有する性質を付与して、音楽データにするものである。フラクタル図形によって生成された樹木の形状や雪の結晶の形状を基に曲を生成する方法が古くから研究されている [1]。我々

は人と人の音声対話における音声の韻律情報からメロディラインを生成する研究を既に行った [2]。このタイプでは、素材としたパターンの性質と生成された音楽の印象が類似しているかどうかよりも、好みの音楽を制作する支援技術としての有用性に重点が置かれる場合が多い。

2つ目のタイプは、可聴化と呼ばれる技術分野であり、音として聴こえない対象物、例えば映像や生体信号などを音響信号や音楽データに変換して聴取可能にすることにより、対象物の特性を聴覚情報に埋め込んで提示し、特性把握の助けにするものである。可聴化では対象物の特性とそれを変換した音の印象との対応関係が重要であり、心地よさや音楽性は副次的である [3][4]。

本研究は1つ目のタイプである。具体的には、国土地理院から大量に提供される山脈の標高データを用いてメロディを生成する。初心者の音楽制作支援を目的とし、音楽

¹ 広島工業大学
Hiroshima Institute of Technology, Hiroshima 731-5193, Japan

² 九州大学大学院人間環境学府
Graduate School of Human-Environment Studies, Kyushu University, Fukuoka 812-0053, Japan

^{a)} y.umemura.im@it-hiroshima.ac.jp

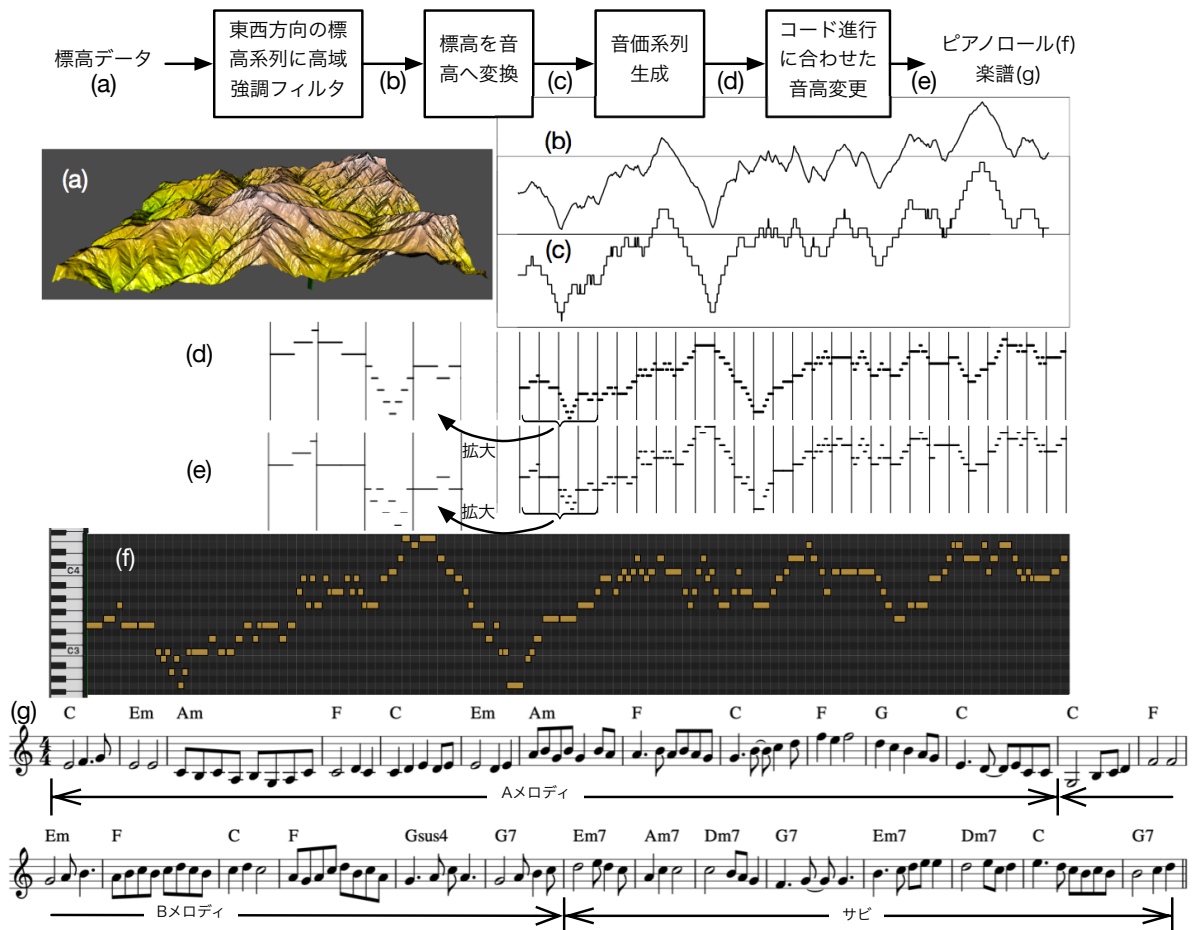


図 1 楽曲生成のブロック図

のジャンルとしてポップス曲の主旋律生成を重点に置く。生成曲の評価として、ユーザーに好まれる音楽を生成できるかが重要である。従来研究において、好みの評価が十分に行われていなかったようである。音楽理論からの逸脱がないかや、Valence-Arousal 平面 (感情価-覚醒度) での位置づけの評価が中心であった (例えば [5])。本研究では、人に好まれるか曲を生成できるかを評価する。本発表は、前回の発表 [6] に対し生成法はほぼ同じで、生成曲の特性の分析と、生成曲の好みの評価実験に重点を置く。

2. 楽曲生成法と支援システムの構想

2.1 生成法

本生成法の基本処理は先の発表 [6] と同じである。処理の流れと各ステップで生成される中間出力を図 1 に示す。主旋律の生成段階での改良点としては、同図のブロック図の最後段に追加した、同音進行を見つけて刺繍音に変更する処理である。

特に、起伏の少ない標高カーブの場合、同音進行が何音符も連続することがある。そこで、同音進行を刺繍音に変更する処理を付加する。刺繍音は、ある音符から次の音符へ音高が 2 度上昇ないし下降し、再び初めの音高へ戻る音高変化である。小節単位で同音進行を検出して刺繍音に変

更する。刺繍音には上昇してから下降するタイプと、下降してから上昇するタイプがある。本アルゴリズムでは刺繍音が出現するたびにタイプを切り替える。小節の終わりで 2 音符の同音進行が生じる可能性および、小節の切り替わりで同音進行が生じる可能性が残っている。

2.2 作曲支援システムの構想

上記生成法により、標高データとコード進行が与えられると主旋律を生成できる。これに様々な機能を付加してシステム化し、初心者用の作曲支援システムとする構想を述べる。いずれも原理的な困難はない。なお、これらの項目のいくつかを今回の主観評価実験用の生成曲作成時に盛り込んでいる。

- (1) 様々なコード進行パターンと伴奏のリズムパターンを、曲中の各セクションごとにメニュー選択できるようにする。
- (2) パーカッションなどの効果音をユーザーがメニュー選択して付け加えられるようにする。
- (3) 曲構成に合うようにフレーズ単位で音域を上下できるようにする。
- (4) 標高データの前後を逆転させるなどの前処理により、後半を盛り上げるなどに対応させる。

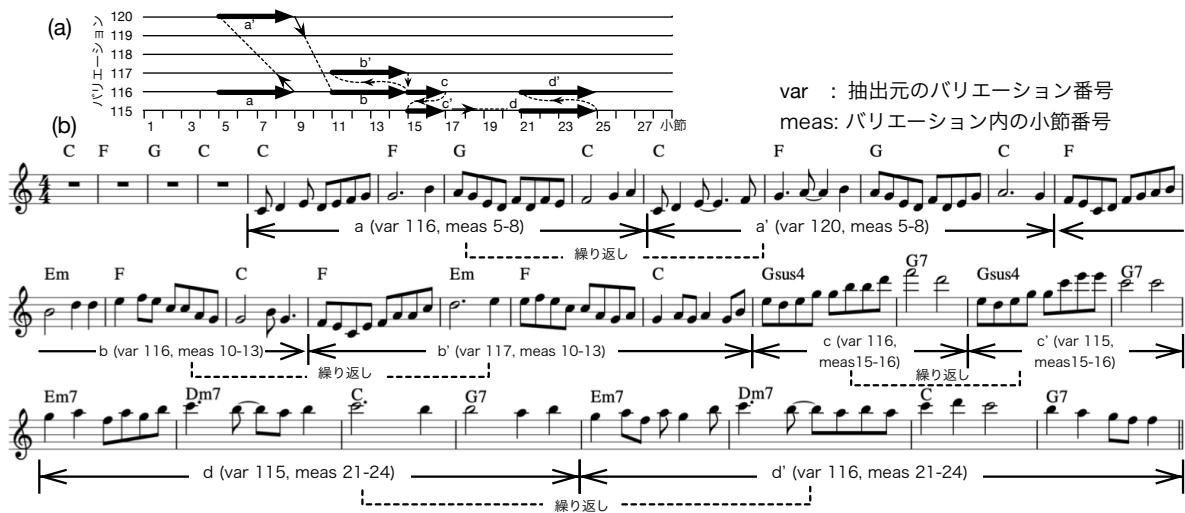


図 2 近傍のバリエーションを使ったフレーズの繰り返し

- (5) 2次元データである標高データから、登山をする人に対して、登山ルート of 標高に基づく曲を生成し、登山の際のBGMを提供する。
- (6) 人間の作曲でフレーズを少し変形させて繰り返すことが行われる。一方、標高データからある東西のラインに沿って主旋律を生成し、その北隣や南隣の東西のラインから別の主旋律を生成すると、両者の同じ程度付近で、類似した主旋律が生成される。それらを使って類似したフレーズを繰り返すことができる。

第6項目の事項に関して、バリエーションという用語を次のように定義する。地形データの1区画に含まれる複数の東西のラインから生成されるそれぞれの曲をバリエーションと称する。

3. 生成曲の評価

3.1 フレーズの繰り返しを伴う曲の作成と評価

先の「作曲支援システムの構想」第6項目のフレーズが変形して繰り返す機能について有用性を調べる。間ノ岳(表2)の標高データから曲を生成し近傍のバリエーションから類似したフレーズを抽出し、元のフレーズの直後に配置して、繰り返しとする。どのバリエーションのどの小節から順に抽出したかを図2(a)に示し、作成した曲の楽譜を同図(b)に示す。なお、楽譜中、d, d'で示す8小節はサビに使用する意図で配置したものの、音域が低いため、1オクターブ上昇させた。先に述べた構想の第3項目に該当する。

以上を基に、作編曲を趣味としている理系大学の在學生と卒業生合計3名に、作成した曲を聴いてもらい、作曲支援システムの中に变化したフレーズが繰り返す機能があれば有用かどうか尋ねたところ、3名とも有用であるとの回答を得た。

3.2 生成曲と楽曲コーパスの統計量の比較

本節では、生成曲の統計量を、楽曲コーパスの統計量と比較することにより、生成曲の音楽性(音楽として自然かどうか)について分析する。

生成曲は北岳を基にして、ハ長調のコード進行に沿って生成した188バリエーションで、主旋律の音符数の合計が20,102音符である。楽曲コーパスは世界の民謡を収録したEssenコーパス[9]中のヨーロッパ曲6,202曲および、コード進行との関係を調べる際には、コード伴奏の付属したコーパスで、英米のフォークソング1,200曲からなるNottinghamコーパス[10]を用いる。

3.2.1 隣接音符間の音程のヒストグラム

旋律によく出現する輪郭パターンとして、順次進行を多く含む滑らかな印象を与える軸型およびアーチ型と、適度な跳躍進行を含みメリハリの効いた印象を与える跳躍充填型がある[11]。輪郭パタンの総合的な傾向を比較するには、大量の旋律におけるすべての隣接音間の音程を集計して、音程の種類ごとの頻度を比較すればよい。Essenコーパスを分析にあたり、文献[12]のp.58の図の方法に準拠した。結果を図3に示す。Essenコーパスでは同音進行が最も多く、2半音下降する進行とほぼ同数であるのに対し、生成曲では同音進行が2半音下降する進行の半数となっている。生成曲の同音進行と2半音上昇する進行の頻度が入れ替わっている点を除くと両者はかなり似たパターンを示している。

3.2.2 音符のピッチクラスのヒストグラム

音階の第1音は主音と呼ばれ、主音からの度数により下屬音、属音等が決められ、それぞれの役割を担う。先の文献[12]p.57の図では、Essenコーパスのヨーロッパ曲について音高の頻度分布を求めている。ここでは機能に着目するため、長調の曲について、ハ長調に移調し、ピッチクラスごとの頻度分布を求める。生成曲も同様にピッチクラス

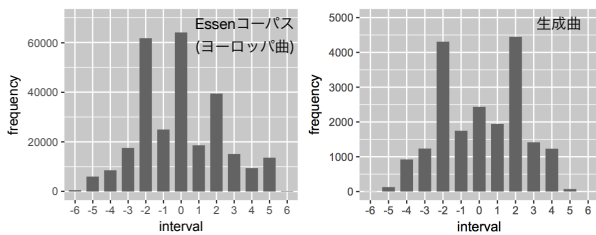


図3 Essen コーパスと生成曲における隣接音符間の音程のヒストグラム

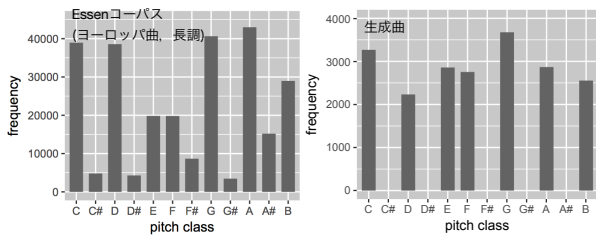


図4 Essen コーパスと生成曲におけるピッチクラスのヒストグラム

ごとの出現頻度を求める。結果を図4に示す。

出現頻度の1位は両者とも属音(G)で、2位も両者とも主音(C)となっている。3位はコーパスでは中音(E)であるのに対し、生成曲では中音(E)と下中音(A)がほぼ同数となっている。このように、類似したパターンを示している。

3.2.3 小節の第1音符の音価と第2音符以降の音価の比

Essen コーパス中のヨーロッパ曲 6,202 曲に含まれる 314,524 音符について、小節の第1音符の音価の平均値と第1音符以外の音価の平均値の比を求めると 1.64 であった。一方、北岳の 188 バリエーション中の 20,342 音符について、小節の第1音符の音価の平均値と第1音符以外の音価の平均値の比を求めると 1.84 であった。音価制御アルゴリズムにより、実存曲の特性に近い値が得られている。

3.2.4 和声音と非和声音の音符数の比

Nottingham Database の 987 曲に含まれる 62,340 音符について和声音と非和声音の音符数の比を求めると 2.9 であった。一方、北岳の 188 バリエーション中の 20,342 音符について和声音と非和声音の音符数の比を求めると 2.0 であった。生成曲の方が非和声音が多く含まれているものの、実存曲の特性に近い値であると思われる。

本節で調べた4つの統計量がすべて、コーパスと生成曲で類似のパターンを示したことから、生成曲の音楽性が高いと考えられる。

3.3 主観評価実験1

本節では地形作曲によって人に好まれる曲を生成できるかどうかについて主観評価実験によって調べる。生成曲は1曲で比較のために用意する人間の作曲による既存曲は5曲である。

3.3.1 実験提示曲

評価者の負担を少なくするため、曲の一部としてサビと思われる部分を切り出して用いる。切り出し位置はフレーズの境界位置である。終わりに1秒間フェードアウトを入れている。主旋律のみの演奏である。演奏用にDTMソフト Logic Pro X を用い、楽器としてギター “12 String Acoustic” を使用する。曲名リストを表1に示す。表の第1列に示す名称は実験結果のグラフで用いる名称に該当する。曲の提示順をランダム化する。

地形作曲による生成曲は、剣岳の標高データ (FG-GML-5437-74-DEM5A) を用いて生成し、先に述べたフレーズの繰り返し方法に基づいて隣接バリエーションからフレーズの繰り返しを生成する。楽譜を図5に示す。

既存曲5曲の選定にあたって、次の点を考慮した。

- 作曲支援システムとして想定するジャンルであるポップス曲から選定する。
- 非常に有名な曲を避ける。評価者が知っている曲の場合、先入観に影響される恐れがある。集計の際、曲を知っている評価者を除外する。
- 前項を満たしつつ、かつ、チャートのある程度上位を占めた曲やテレビ主題歌に使われた曲を選定する。

3.3.2 実験方法

理系大学の大学1年生から大学院1年生までの一般評価者62名に曲を評価をしてもらった。曲の提示順をランダムに設定した6曲が順に演奏され、再度、同じ順で2回目が演奏される。曲の演奏順はランダムに設定した。曲番号などのナレーションも合わせて全体で実験刺激の再生時間が4分41秒であり、スピーカーによる受聴とした。評価の選択肢は次の5つとした。[1. 好まない], [2. やや好まない], [3. どちらとも言えない], [4. やや好む], [5. 好む]

評価者に知っている曲かどうかを申告してもらい、1曲でも知っている曲がある評価者を集計の段階で除外した。上記の人数62名は除外後の人数である。

3.3.3 実験結果および考察

生成曲と既存曲の評価値から分散分析を行った結果、曲を要因として有意差が認められた。次に生成曲と各既存曲を多重比較するための検定方法として、多対一比較が可能な Dunnett の方法 [13] によって検定した結果、有意水準5%にて曲1と曲3に比べて生成曲が有意に好まれるとの結果を得た。図6に既存曲5曲と生成曲の評価値の差をプロットする。左ほど、生成曲が好まれることを示す。エラーバーは95%信頼区間を意味する。

実験の結果、地形作曲によって生成された主旋律は、人間の作曲者による既存曲の主旋律よりも、好まれる旋律となる場合があることが示された。

曲が人々に好まれるかどうかは、主旋律だけでなく様々な要素によって影響を受ける。しかし、その中でも主旋律は重要な要素の1つである。また、初心者の音楽制作支援

表 1 実験提示曲の曲名リスト

ID	曲名	作曲者	発売	ジャンル	時間
生成曲	(劔岳より生成)				18 秒
曲 1	Strangers by night	Dieter Bohlen	1986	ポップス	22 秒
曲 2	始まりの音色	浜口史郎	2011	主題歌	19 秒
曲 3	Baby I love your way	Peter Frampton	1976	ポップス	19 秒
曲 4	緑色の記憶	吉森信	2010	主題歌	19 秒
曲 5	Beautiful Life	Jonas Berggren	1995	ポップス	16 秒



図 5 実験 1 に用いた劔岳の標高データからの生成曲

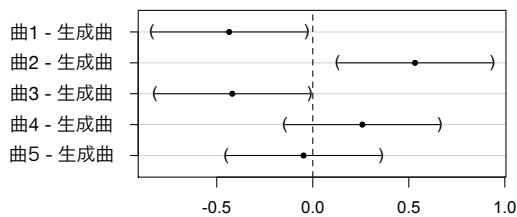


図 6 既存曲と生成曲の差に関する分散分析。エラーバーは 95%信頼区間を示す

を想定すると、主旋律以外の要素は部品をメニューから選んで配置するシステムによって支援できると思われる。現行の GTM ソフトでは、例えばドラム伴奏に関して、膨大なパターンからメニュー選択して曲の中に配置できるようになっている。それに対して主旋律は自由度が大きく、パターンをメニューで選ぶタイプの支援に適さない。したがって、人に好まれる主旋律を生成できるのは有用な技術であると考えられる。

3.4 主観評価実験 2

実験 1 で用いた生成曲は 1 曲である。そこで様々な山から生成しても人に好まれる曲が生成できるかどうかを調べるため、実験 1 で用いた生成曲およびそれ以外の 6 山に基づく 6 曲の計 7 曲を用いて主観評価実験を行う。

3.4.1 実験提示曲

7 曲の生成方法について、まず、実験 1 との違いを大まかに述べる。与えるコード進行は J-pop を想定したもので、各山ともほぼ共通とした。各山から生成される複数バリエーション中の 1 バリエーションを実験者が主観的に選定して提示曲とした。実験 1 と異なりフレーズの繰り返し処理を行わない。演奏に際し、主旋律にコード伴奏を付けた。図 7 に劔岳から生成した楽譜を示す。

次に詳細を述べる。用いる 7 曲の仕様を表 2 に示す。このうち、北岳は先の図 1 中に記載した楽譜の曲である。他の 5 曲の楽譜は割愛する。標高データの前処理として、櫛形山のみ東西を逆転させて上り調子にした (前記、構想の第 4 項目に対応)。

コード進行の与え方は、代表的なコード進行の掲載された書籍 [8] から選択した。具体的には、A メロディに「ちょっと悲しげなコード進行 C-Em-Am-F-C-F-G-C」、B メロディに「期待をもたせるコード進行 C-F-Em-F-C-F-Gsus4-G」、サビに「安定したコード進行 Em7-Am7-Dm7-G7-Em7-Dm7-C-G7」を採用した。コード伴奏の演奏法として、A メロディ、B メロディにはアルペジオ奏法、サビにはバックギン奏法を採用した。イントロは 7 山とも同じコード進行 C, F, G, C からなる 4 小節のコード伴奏を用いた。曲の最後に再び、同じイントロが演奏され、演奏の最後 4 秒間でフェードアウトする。

各セクションの構成の詳細を述べる。乗鞍岳以外の 6 曲は A メロディ、B メロディ、サビの順である。乗鞍岳のみ、主旋律の後半がゆっくり下降するパターンのため、サビの後に A メロディを付け加えた。また、主旋律のパターンに応じてバリエーション内の各セクションの小節数を設定した。DTM ソフト Logic Pro X を用いてテンポ 160 で演奏する。パートに割り当てる音源として、主旋律に管楽器 "Funk Horn Section"、コード伴奏にピアノ "Steinway Grand Piano" とした。

表 2 実験提示曲の曲名リスト。小節数はイントロを除いた値である

ID	名称	ファイル名 (.zip)	小節
-	劔岳 (基準曲)	FG-GML-5437-74-DEM5A	20
1	北岳	FG-GML-5338-41-DEM5A	28
2	奥穂高岳	FG-GML-5437-35-DEM5A	20
3	間ノ岳	FG-GML-5338-41-DEM5A	24
4	櫛形山	FG-GML-5338-32-DEM5A	24
5	乗鞍岳	FG-GML-5437-14-DEM5A	24
6	赤石岳	FG-GML-5338-11-DEM5A	24

3.4.2 実験方法

評価者に関し、実験 1 の 62 名の中から地形作曲を好む評価者を、地形作曲の評価値がそれ以外の 6 曲の評価値の平均よりも 1 以上高いという条件で抽出した。その結果、11 名が抽出された。これら 11 名の評価値を調べたところ、地形作曲の評価値は 4 か 5 のいずれかであり、4 が 9 名、5



図 7 実験 2 に用いた剣岳の標高データからの生成曲

が 2 名であった。

これらの評価者が、実験 1 で用いた曲を評価の基準として、他の 6 曲それぞれについて、基準曲より好むか否かを評価した。曲の提示順は、第 1 セットとして、基準曲、曲 1, ..., 曲 6 の順に演奏され、第 2 セットとして、基準曲、曲 1, 曲 2, 基準曲、曲 3, 曲 4, 基準曲、曲 5, 曲 6 の順に演奏される。比較対象曲 6 曲の提示順を評価者ごとにランダム化した。評価の選択肢は次の 5 つである。[1. 基準曲より好まない], [2. 基準曲よりやや好まない], [3. 基準曲と同じ], [4. 基準曲よりやや好む], [5. 基準曲より好む]

ヘッドホン受聴であり、評価者の判断で聴き直しも可能とした。一度も聴き直しを行わなかった場合、曲番号などのナレーションを含めた全体の再生時間は 13 分 14 秒であった。

3.4.3 実験結果および考察

6 曲それぞれに対する 11 名の評価者の評価値の平均値と $\pm 1\sigma$ の範囲をエラーバーとして付加したグラフを図 8 に示す。評価値 3 が北岳と同等という評価のため、縦軸が 3 の位置に水平線を付加する。

実験 1 で北岳を「好ましい」と評価した 11 名に対し、実験 2 を実施した。実験 2 では北岳以外の 6 山に基づいた生成曲を対象とし、北岳を基準値 3 として 1 (好ましくない) から 5 (好ましい) の 5 段階での主観評価を行った。評価対象とした 6 曲と基準値との間に大きな違いはなかった。

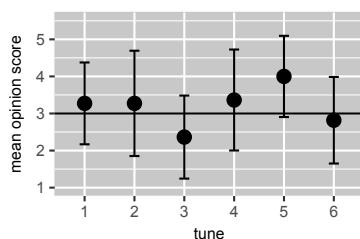


図 8 比較対象曲 6 曲の平均オピニオン評点。エラーバーは標準偏差の範囲を示す

4. まとめ

自然界に存在するパターンから曲を生成する技術として、大量かつ容易に入手できる地形の標高データを使って主旋律を生成する技術を開発した。生成曲の音楽性に関する

客観評価と、生成曲が人に好まれるかに関する主観評価を行った。まず、音楽性について、隣接音符間の音程の頻度分布などの統計量を楽曲コーパスと比較することによって確認した。次に、人間の作曲による既存曲 5 曲と生成曲 1 曲を使って、好みの程度を主観評価する実験を行い、有意差検定により、生成曲が既存曲のうちの 2 曲よりも有意に好まれることを確認した。また、いくつかの地形から生成しても、同様に好まれる曲を生成できるかを主観評価実験によって確認した。

謝辞 本研究は、広島工業大学梅村研究室の卒業研究として検討してきたテーマである。研究室の卒業生、福本亮太君、伊達彩斗君、松崎裕佑君に深謝する。

参考文献

- [1] Gerhard Nierhaus: *Chaos and Self-Similarity. Algorithmic Composition*, Springer(2009).
- [2] 大野拳汰, 山本樹, 佐々並賢志, 梅村祥之: 対話音声の韻律情報を用いたメロディ生成の試み, 第 110 回情報処理学会音楽情報科学研究会 (2016).
- [3] J. Pellerier: Perceptually Motivated Sonification of Moving Images, *Proceedings of the 2009 International Computer Music Conference*, pp.207-210(2009).
- [4] 宮崎敦子, 田中文久, 井出祐昭, 高橋信, 川島隆太: 脳の奏でる音楽とコミュニケーションの関係性について, 日本音響学会研究発表会講演論文集 (2013).
- [5] 深山寛, 中妻啓, 酒向慎司, 西本卓也, 小野順貴, 嵯峨山茂樹: 音楽要素の分解再構成に基づく日本語歌詞からの旋律自動作曲, 情報処理学会論文誌, Vol.54, No.5, pp.1709-1720(2013).
- [6] 梅村祥之, 伊達 彩斗: 地図標高データを用いたメロディ生成の試み, 第 115 回音楽情報科学研究会 (2017).
- [7] 国土地理院: 基盤地図情報ダウンロードサービス (online), 入手先 (<https://fgd.gsi.go.jp/download/menu.php>).
- [8] 石沢功治: [ギター・マガジン] すぐ歌えるコード進行ネタ帳 選んでつなげて 5 万曲!, リットーミュージック (2008).
- [9] <http://www.esac-data.org>
- [10] <https://ifdo.ca/seymour/nottingham/nottingham.html>
- [11] R. Synder: Music and memory, *The MIT Press*(2000) 須藤貢明, 杵鞭広美 (訳): 11 メロディ, In 音楽と記憶, 音楽之友社, pp. 162-165 (2003).
- [12] D. Temperley: Music and Probability, *The MIT Press*(2007).
- [13] C. Dunnett: A multiple comparison procedure for comparing several treatments with a control, *Journal of the American Statistical Association*, Vol. 50, pp. 1096-1121(1955).