

印象を考慮した音楽生成手法の検討

渡邊 祥太郎^{†1,a)} 井上 勝文^{†1,b)} 吉岡 理文^{†1,c)}

概要: 音楽生成においては、聴取者の興味を惹く多様な音楽の生成が求められている。特に、音楽を聴取した際の印象を加味できる音楽生成モデルが求められている。本研究では、音楽の印象を表す印象ベクトルとメロディを入力とし、印象ベクトルに応じた和音進行とテンポを出力とする、Long Short Term Memory(LSTM)を用いたモデルを提案する。また本研究では、提案モデルを用いて生成した和音進行とテンポが、入力時に指定した印象ベクトルに沿っているか定性調査する。

Investigation of Music Generation Based on Impression

1. はじめに

音楽は一般的にハーモニー・リズム・メロディの三要素によって構成される。これに基づいて音楽を作曲する際には、ハーモニーを表現するために和音進行を、リズムを表現するためにテンポやドラムパターンを、またメロディをそれぞれ作成する。このような作曲を行う中でも、メロディに基づいた和音進行を作成することは、重要なプロセスの一つとして挙げられる。しかし、これを行うためには音楽理論や和音の性質への理解が必要であり、このような知識を持っていない者にとっては困難な場合がある。この問題を克服するための手法として、メロディに基づく和音進行の自動生成手法 [1-3] や、メロディに基づくドラムパターンの自動生成手法 [4] が提案されている。

しかし文献 [1-4] においては、「どのような印象を聴取者に与える楽曲を作曲したいか」という作曲者の意図が加味されていないという問題がある。心理学的見地において、作曲過程ではどのような楽曲を作曲したいか想起するフェーズが存在すると指摘されている [5] ため、自動生成される楽曲についても作曲者の意図が加味される必要があると考える。ここで、音楽を構成する三要素のうち、ハーモニーとリズムに関連する和音進行やテンポによって、音楽の印象は変化することが知られている [6, 7]。これらを

加味することで、作曲者の意図を考慮した音楽を作成できると予想される。なお、アルペジオに代表されるような和音の演奏方法や、音楽の調によっても作曲者の意図を考慮できると考えられるが、問題の単純化のため本研究においては和音の種類やその進行、また特定の調に限定した場合のみについて検証を行う。

本研究では、あるメロディに持たせたい印象に応じた和音進行とテンポを自動生成できるか検討する。具体的には、時系列データの学習に適した LSTM(Long Short-term Memory) [8] を用いて、音楽の印象を表す印象ベクトルを付与したメロディを入力し、入力した印象を加味した和音進行とテンポを出力する LSTM ベースの Deep Neural Network(DNN) モデルを提案する。本研究では、印象ベクトルには「快度」「覚醒度」という 2 種類の印象観点によって表される、2次元のベクトルを用いる。

ここで本論文で扱う音楽専門用語について表 1 にまとめる [9]。なお本論文ではドイツ音名を使用し、和音名については、C を主音とする長三和音を C、A を主音とする短三和音を Am などと表記する。また、和音の種類として長三和音を maj、短三和音を min などと表記する。

2. 関連研究

本節では関連研究について述べる。2.1 節ではメロディに基づく和音進行の自動生成研究について述べる。また 2.2 節では音楽聴取時における印象に関する研究について述べる。

^{†1} 現在、大阪府立大学 大学院工学研究科
Presently with Graduate School of Engineering, Osaka Prefecture University

^{a)} watanabe_s@sig.cs.osakafu-u.ac.jp

^{b)} inoue@cs.osakafu-u.ac.jp

^{c)} yoshioka@cs.osakafu-u.ac.jp

表 1: 音楽専門用語
Table 1 Technical term of music.

用語	意味
音階	一定の秩序で配列された音の階段. この様子を図 1 に示す.
音度	音階を構成する各音. 主音は I 度と呼ばれ, 以下順次, 上方に向かって II 度・III 度… と数えられる. この様子を図 2 に示す.
和音	一定の秩序によって組み合わせられたいくつかの音の集団.
三和音	ある音の上に III 度と V 度の音を積み重ねた和音. 構成する 3 つの構成音相互の音程関連によって図 3 に示す 3 種類に分けられる.

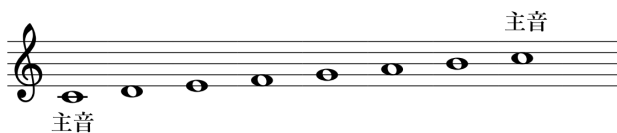


図 1: 主音を C とする音階

Fig. 1 Scale in C major.

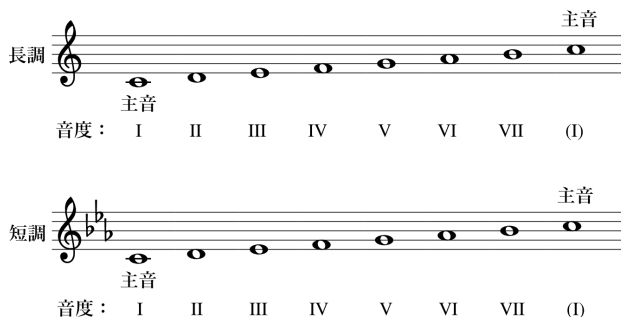


図 2: 主音を C とする長調・短調における音度

Fig. 2 Scale degree in C major and minor.

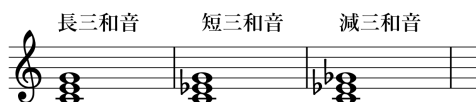


図 3: 主音を C とする三和音の種類

Fig. 3 Types of triads in C major.

2.1 メロディに基づく和音進行の自動生成

あるメロディに対して, これに沿った和音進行を生成することは作曲者にとって重要なプロセスの一つである. これを自動化するために, 隠れマルコフモデル (HMM) を用いた手法 [1,2] が検討されている. しかし, HMM においては時系列データの要素における長期的な依存関係を学習できないという問題点がある. そこでこの問題点を解決するために, Lim らは LSTM ベースの DNN モデルを用いた手法について検討している [3]. しかし [3] において入力する情報は, メロディを構成している各音の構成とその小節内における割合のみであり, メロディの各音が鳴る順番は入力の情報として考慮されていない. また [3] において出力する情報は和音進行のみであり, テンポについては考慮されていない. メロディの鳴る順番やテンポは音楽の要素として重要であるため, 本研究ではこれを考慮した LSTM

ベースのモデルを提案する.

また, これらの手法は聴取者が生成された音楽を聴いた際に受ける印象についても考慮していないため, 本研究では印象情報を前述のモデルに組み込み, 印象を加味した和音進行を生成する手法を提案する.

2.2 音楽聴取時における印象

三和音が聴取者に与える印象について, [10] では「快さ」「あわただしさ」「安定性」という印象観点をを用いて検討している. この研究においては和音の転回形などを含め計 7 種類の三和音を聴取対象としているが, 特に根音を C とする長三和音と短三和音についてのみ着目すると, 長三和音は短三和音に比べ「快」「ゆったりとした」「安定した」印象を持つとしている.

また Ritossa ら [11] は, Russell の円環モデル [12] によって分類される各音楽が聴取者に与える影響について調査している. ここで, Russell による円環モデルは, 感情を「快 (Pleasant)–不快 (Unpleasant)」と「覚醒 (Arousing)–沈静 (Sleepy)」という 2 つの軸によって平面上に配することができるとしている. Ritossa らは Russell の円環モデルにおける 4 つの象限それぞれの感情が誘起される音楽を予備実験によって選出し, これらの音楽があらゆる印象を持つ音楽であるとした上で, 改めてこれらの音楽が複数の印象観点に対しどのような評価値を持つかをアンケート調査している. この際用いた観点は, 4 つの象限それぞれにおいて 2 つずつ用意された計 8 つの形容詞および, 「覚醒 (Arousal)」「快 (Pleasantness)」「好み (Liking)」「親密度 (Familiarity)」である. また 8 つの形容詞に対する評定値を従属変数, 「覚醒」「快」「好み」「親密度」の評定値を説明変数とする重回帰分析を行っている. その結果, 「覚醒」と「快」によって各形容詞の評定を概ね説明できることを示している.

本研究では [10, 11] に共通する音楽の印象観点として, 「快」に関する尺度を示す「快度」, および「覚醒」に関する尺度を示す「覚醒度」という観点をを用いる.

3. 提案手法

本節では提案手法について述べる. 提案手法では, 指定した印象を加味した和音進行とテンポを生成できるように, 音楽の印象を表す印象ベクトルをモデルの入力に用いる. このモデル構造については 3.1 節で, モデルの入力に

用いる印象ベクトルについては3.2節で、モデルの入力に用いるデータセットの前処理については3.3節で、モデルの学習については3.4節で述べる。

3.1 モデル構造

本研究では、Limらの手法と異なりメロディで鳴らす各音の構成とその割合のみでなく、その鳴らされる順番やタイミングを考慮し、また時系列データを扱うため、Encoder-Decoderモデルのひとつであるseq2seq [13]を基にモデル構築を行う。本研究で作成したモデル構造を図4に示す。

提案モデルには、音楽の印象を表すベクトルを付与したメロディを入力する。また、和音進行やテンポによって音楽の印象は変化するという知見 [6,7] に基づいて、和音進行とテンポを出力する。EncoderとDecoderのシーケンス長はそれぞれ、入力に用いるメロディデータの時系列方向の長さ N と、出力に用いる和音進行データの時系列方向の長さ M とする。なお、メロディと和音進行が時系列データであるのに対し、印象ベクトルとテンポデータは非時系列データである。これらのデータを同時に提案モデルの入出力として用いるために、印象ベクトルをEncoderにおける全シーケンスで入力し、テンポデータをDecoderにおける全シーケンスで出力する。

3.2 印象ベクトル

本研究では音楽の印象に関して、[11]を参考に「快度」「覚醒度」という観点を用いる。印象ベクトルとして、「快度」「覚醒度」それぞれについて0から1の連続値で表現される評価値を持つ2次元のベクトルを用いる。具体的には、「快度」の評価値は「快」という印象を持つ音楽ほど1、「不快」という印象を持つ音楽ほど0に近づく。学習に用いる印象ベクトルの計算は以下のように行う。まず、ある一定の小節数を持つ和音進行を複数用意する。なおこれらの和音進行は、それぞれが異なるテンポであるものとする。これらに対して、評価者は「快度」「覚醒度」それぞれにおける評価値を与える。例えば、「快度」の項目では「快:5, やや快:4, どちらでもない:3, やや不快:2, 不快:1」の5段階である。この評価値に対する各評価者間の平均値をとり、0から1の値に正規化したものを印象ベクトルの要素とする。同様に、「覚醒度」においても「覚醒:5—沈静:1」として0から1の連続値を印象ベクトルの要素として用いる。なおこの際、評価者における評価値の分散は考慮していない。また印象ベクトルの表記について、本研究においては、「快度」が x_1 かつ「覚醒度」が x_2 である印象ベクトルを $[x_1, x_2]$ と表すこととする。

3.3 前処理

モデルの入力に用いるメロディデータセットと、出力に

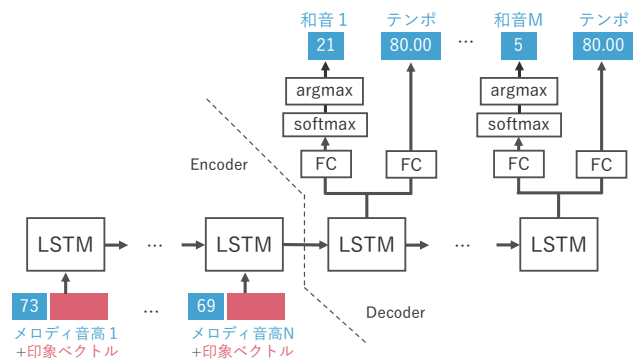


図4: 提案手法のモデル構造

Fig. 4 Overview of proposed model.

用いる和音進行・テンポデータセットに対して、前処理を行う。

まずメロディデータセットから、8小節の長さを持つメロディを s 個選出する。これらのメロディにおいて *Note Number* を符点8分音符間隔で取得し、メロディデータとする。1小節の長さは符点8分音符を24個並べた長さと同じいため、各メロディデータにおける時系列方向の長さは $24 \times 8 = 192$ となる。ここで *Note Number* は、周波数 $f[\text{Hz}]$ に対し

$$\text{NoteNumber} = 12 \log_2 \frac{f}{440} + 69$$

で定義される値とする。*Note Number* と周波数 $f[\text{Hz}]$ の関係を表2に示す。

また和音の種類は非常に多く、すべての和音の種類を分類の対象とした場合、各種類における十分なサンプルがテンポデータセット内に含まれない可能性がある。本研究では問題の簡単化のために、まず和音の置換によって用いる種類を減らす。和音進行データセット内に含まれる和音のうち、根音12種類に対する長三和音・短三和音の計24種類以外の和音は、根音が同じ長三和音・短三和音いずれかに置換する。この置換は表3に従って行う。

次に、和音の鳴らされる長さによる印象への影響を避けるため、和音を調整する。具体的には、各和音が全て全音符によって構成されるように、各小節の1拍目で鳴らされている和音を各小節における和音として置換する。

さらに楽曲の調による印象への影響を避けるため、各楽曲の調を統一する。和音進行データセット内の各楽曲で示されている調に従って、各楽曲が長調であれば長調に、短調であれば短調にそれぞれ移調する。これらの調はいずれも白鍵のみによって音階が構成されているものである。

以上の処理を、和音進行データセットから抽出した s 曲に対して行い、これを和音進行データとする。また同時に、各和音進行を抽出した楽曲が持つテンポを t 個抽出し、これらを各和音進行に対するテンポデータとする。

表 2: Note Number と周波数の関係
Table 2 Relationship between Note Number and frequency.

Note Number	周波数 [Hz]	音名
0	8.2	C
1	8.7	C#
2	9.2	D
⋮	⋮	⋮
68	415.3	G#
69	440.0	A
70	466.2	A#
⋮	⋮	⋮
125	11175.3	F
126	11839.8	F#
127	12543.9	G

3.4 学習

モデルの学習においては 3.1 節で記した通り, Encoder にはメロディデータと印象ベクトルが入力し, Decoder には和音進行・テンポデータが出力する. また, それぞれの印象ベクトルを算出する際に用いた和音進行とテンポを, 印象ベクトルに対応する和音進行・テンポデータの Ground Truth とする. この際和音進行データは, 12 種類の根音に対する長三和音・短三和音の計 24 種類の和音をクラスとする, 24 次元の One-hot ベクトルとする.

モデルの学習時には, 和音についての損失関数として

$$L_{CE} = - \sum_{i=1}^c b_i \log p_i \quad (1)$$

で表される値を用いる. このとき c は出力される和音の数を, b_i は i 番目に出力される和音の種類が Ground Truth と等しい場合にのみ 1 をとるバイナリ値を, p_i は i 番目の出力における予測確率を表す. またテンポについては損失関数として

$$L_{MSE} = (y - \hat{y})^2 \quad (2)$$

で表される値を用いる. このとき y は出力されるテンポの予測値を, \hat{y} は出力されるテンポに対する Ground Truth を表す. なお学習済みのモデルにおいて入力されたメロディと印象ベクトルに対してテンポの値を求める際には, Decoder の全シーケンスから得られたテンポデータの値の平均を各入力に対するテンポの値として扱う. そしてこれらの損失関数で表される値を最小化するように学習を行う. 最終損失関数は L_{CE} , L_{MSE} を用いて

$$L = L_{CE} + L_{MSE} \quad (3)$$

で表される値を用いる.

4. 実験

本節では提案手法を用いて出力される和音進行とテンポ

表 3: 和音の種類の変換表
Table 3 Chord type conversion table.

置換前	置換後
maj, 5, 7, M7, aug, add9, sus2, sus4, 7sus4, 6, 9, 11, 13	maj
min, m7, m7-5, mM7, dim, dim7, m6, m9, m11, m13	min

が, 入力に用いた印象ベクトルに沿っているかを検証する. 本実験ではメロディのデータセットとして The NES Music Database [14] のパルス波トラックを用いた. また, 和音進行とテンポのデータセットとして The McGill Billboard Project [15] によって提供されるデータセット (以下, Billboard Dataset) を用いた. これは 1958 年から 1991 年のビルボードチャートにおける一部の楽曲について, 和音進行や調などを記したデータセットである. ここでメロディデータセットと和音進行・テンポデータセットを異なるものに指定した理由は, 作曲者が聴取者に与えたい印象を考慮して作曲した結果, 敢えてメロディに沿わない和音進行やテンポを選択する可能性があると考えたためである.

印象ベクトルの算出については 4.1 節で, 図 4 に示したモデルの学習とテストについてはそれぞれ 4.2 節で述べる. また生成した音楽に対する主観評価については 4.3 節で述べる.

4.1 印象ベクトルの算出

本実験では, Billboard Dataset からランダムに抽出した $s = 20$ 曲の和音進行・テンポデータに対し, 評価者 11 名が「快度」「覚醒度」それぞれについて 1 から 5 の 5 段階評価を回答し, 3.2 節に示した方法で印象ベクトルの算出をした. 算出に用いた 20 曲の和音進行・テンポと対応する印象ベクトルの一覧を表 4 に示す. また表 4 に示した 20 曲分の「快度」「覚醒度」の分散はそれぞれ 0.0191, 0.0391 であった. これらの分散値から, 「快度」に比べ, 「覚醒度」に関してより多様な音楽がこの 20 曲中に含まれていると考えられる.

4.2 学習・テスト

図 4 に示したモデルを学習する際, 和音進行データとメロディデータのサイズに従って, モデル構造における Encoder と Decoder のシーケンス長 N , M はそれぞれ $N=192$, $M=8$ とした.

また, The NES Music Database からテスト用のメロディ 1 曲を抽出した. この 1 曲のメロディと, 「快度」「覚醒度」それぞれに 0, 0.5, 1 を用いた 9 種類の印象ベクトルを学習済みのモデルに入力した. ここで, 出力された和音進行に入力したメロディを加え, 同時に出力されたテンポで速度を定めた 9 種類の音楽を作成した.

表 4: 20 曲の和音進行とテンポに対する印象ベクトルの値
Table 4 Values of impression vector for chord progressions and tempos.

和音進行								テンポ	印象ベクトル	
和音 1	和音 2	和音 3	和音 4	和音 5	和音 6	和音 7	和音 8		快度	覚醒度
Em	Dm	G	Dm	Em	Dm	G	F	155.94	0.350	0.500
F	G	G	C	C	Am	Am	F	158.89	0.500	0.650
Em	F	G	C	Em	F	F	C	96.6	0.625	0.400
C	F	Fm	D	G	G	C	C	126.97	0.300	0.250
C	C	C	C	C	C	C	C	116.71	0.475	0.275
Cm	Cm	C	C	C	Cm	A#	A#	143.27	0.325	0.575
C	C	F	F	C	C	Am	Am	168.24	0.725	0.625
G	F	F	Am	Am	F	F	G	103.84	0.625	0.300
C	F	G	E	Am	F	G	C	113.12	0.500	0.550
C	G#	C	A#	C	G#	Fm	C#	118.45	0.300	0.600
F	F	C	G	F	F	C	G	181.19	0.675	0.775
F	C	F	C	F	C	F	C	187.87	0.500	0.600
G	C	C	F	G	G	C	F	65.75	0.250	0.200
Dm	C	F	Dm	Dm	C	Dm	C	100.72	0.425	0.225
C	G#	Dm	E	Am	D	Dm	G	50.85	0.250	0.125
C	C	F	F	C	C	F	F	99.77	0.525	0.300
F	G	F	G	F	G	Am	Am	107.45	0.525	0.350
C	G	F	G	C	G	F	G	148.57	0.600	0.875
G	G	Em	Em	G	G	Em	Em	126.24	0.525	0.450
F	C	G	G	F	F	C	C	125.18	0.550	0.525

入力した印象ベクトルに対する出力されたテンポを表 5 に示す。表 5 より、「快度」と「覚醒度」のいずれにおいても、この値の高低にしたがって、生成された音楽におけるテンポが概ね増減していることが見受けられる。即ち、テンポの値の高さは「快度」と「覚醒度」の値の高さに影響を及ぼす傾向があると考えられる。

テスト時の入力に用いたメロディを図 5 に示す。また印象ベクトル [0, 0], [0, 1], [0.5, 0.5], [1, 0], [1, 1] を入力に用いて作成した音楽を図 6 に示す。なお各楽譜の左部には入力に用いた印象ベクトルを、各和音の上部には和音名を示した。印象ベクトル [0, 0] と [0, 1] によって生成された音楽の比較や、印象ベクトル [1, 0] と [1, 1] によって生成された音楽の比較によると、覚醒度の値を高く設定することによって、生成される和音進行に含まれる長三和音の割合が高くなる傾向が見受けられる。この傾向を出力された全ての和音進行について考察するため、生成された和音に含まれる長三和音の割合 [%] を表 6 に示す。表 6 より、「覚醒度」の値が 1 である場合、作成した音楽の和音進行には長三和音のみが含まれており、長三和音が含まれる割合の高さが「覚醒度」の高さに影響していると考えられる。しかし、岩宮らによると、長三和音は短三和音に比べ「ゆったりとした」印象を聴取者に与えるとしている [10]。[10] における「ゆったりとした」印象が本研究における「覚醒度」の低さに相当すると仮定すると、表 6 に示した結果は [10] に沿っていないと考えられる。しかし、[10] においては和

表 5: 入力された印象ベクトルに対して出力されたテンポ
Table 5 Values of output tempos with input impression vectors.

印象ベクトルの要素	覚醒度=0	覚醒度=0.5	覚醒度=1
快度=0	45.34	137.84	123.58
快度=0.5	67.94	120.39	151.10
快度=1	113.42	169.86	173.94

表 6: 入力された印象ベクトルに対して出力された和音進行に含まれる長三和音の割合 [%]
Table 6 The percentage of major triads in output chord progressions with input impression vectors.

印象ベクトルの要素	覚醒度=0	覚醒度=0.5	覚醒度=1
快度=0	50.0	37.5	100.0
快度=0.5	37.5	87.5	100.0
快度=1	75.0	75.0	100.0

音が持つ印象についてのみ検討しており、本研究では和音進行が持つ印象について検討しているため、和音の連なり方によってその和音進行全体の印象は変化すると考えられる。

4.3 生成した音楽に対する主観評価

生成した 9 曲の音楽が入力した印象ベクトル通りの印象を聴取者に与えるか、4 名の評価者による主観評価で検証を行った。この検証では生成した 9 曲の音楽に対してシェッ



図 5: テスト時に入力したメロディ

Fig. 5 Testing input melody.

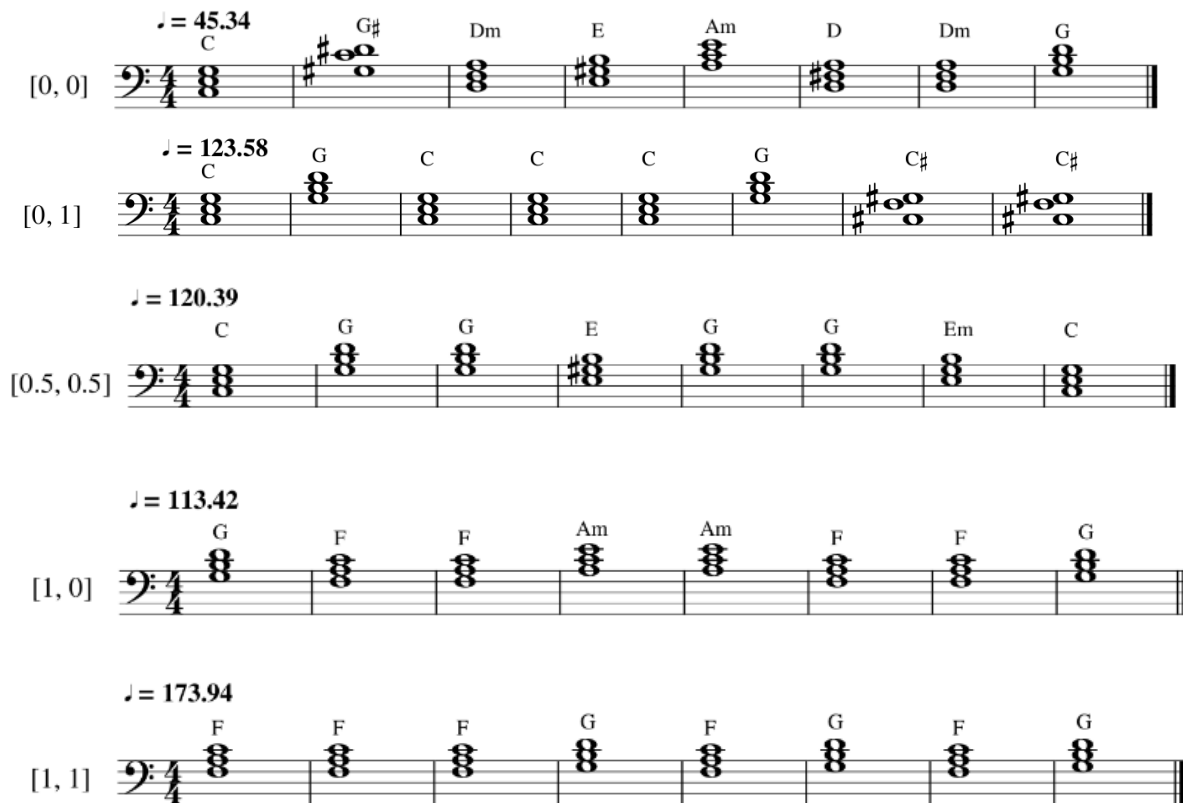


図 6: 印象ベクトルを入力して生成された音楽の例

Fig. 6 Examples of generated music with input impression vectors.

フェの原法 [16] を以下のように行った。まず、作成した音楽から 2 曲を抽出し、これに対して評価者が 1 対 1 で比較した。音楽 A と音楽 B を連続で聴き、音楽 B は音楽 A に比べてどの程度異なるかを評価した。例えば、「快度」の項目では「快:+2, やや快:+1, 同等:0, やや不快:-1, 不快:-2」の 5 段階である。これを全ての対について行い、また「覚醒度」についても同様に行った。入力した印象ベクトルごとに評価値を回帰分析して偏回帰係数を算出し、これらを各音楽についての相対的評価値とした。この相対的評価値を図 7 に示す。

図 7 より、「快度」について、印象ベクトルに [1, 1] を用いて生成した音楽は最も快寄りに配されており、この音楽が入力した印象ベクトルの「快度」を概ね反映していることが分かる。また「覚醒度」について、印象ベクトルに [0, 0] を用いて生成した音楽は最も沈静寄りに配されており、この音楽が入力した印象ベクトルの「覚醒度」を概ね反映していることが分かる。これは表 5 に示されているよ

うに、「快度」や「覚醒度」の高低にしたがって、生成された音楽におけるテンポが概ね増減していることに起因すると思われる。

しかし、印象ベクトル [0.5, 1] を入力して作成した音楽が最も「不快」寄りに配され、印象ベクトル [1, 0.5] を入力して作成した音楽が最も「覚醒」寄りに配されているなど、入力した印象ベクトルに沿わない音楽が生成される例も見受けられた。これは印象ベクトルにおけるいずれかの要素の値を 0.5 に設定した音楽が、各聴取者に様々な印象を与えたことに起因すると思われる。

また印象ベクトルの要素に 0.5 を設定していない例のうち、印象ベクトル [1, 0] を入力して作成した音楽が「覚醒」寄りに配されているという例も見受けられた。これは入力した「覚醒度」の値が低い場合においても、同時に入力した「快度」の値が高い場合には、生成された音楽は「覚醒」であると評価される傾向があると考えられる。これは「快」「覚醒」という 2 つの印象観点によって音楽の印象を表現

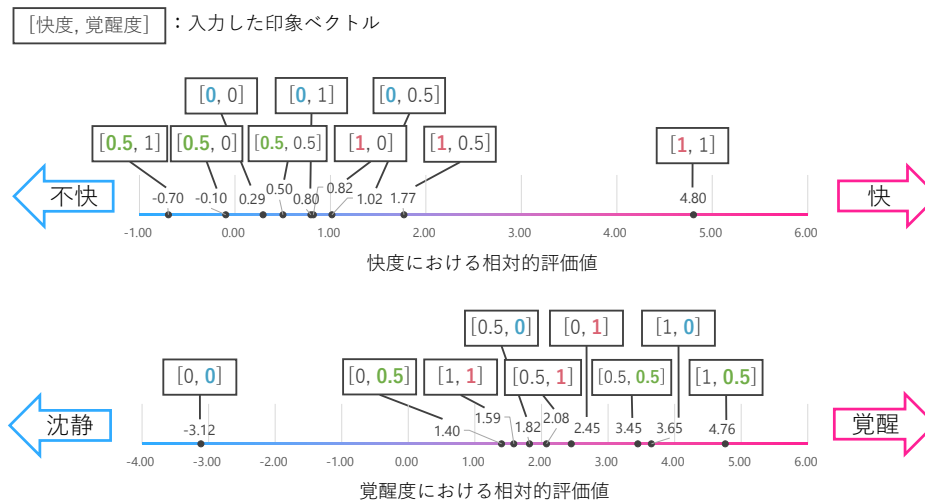


図 7: 入力した印象ベクトルと作成した音楽の相対的評価値

Fig. 7 Relative evaluation values of generated music from input impression vectors.

しきれず、「快感」が「覚醒度」に影響を及ぼしたことに起因すると考えられる。

上述の問題点を回避するためには、より多様な印象観点の導入により、観点間の影響をより少なくする必要があると考えられる。また評価者の音楽の好みや音楽経験によって定性評価の結果は変動すると考えられるため、より豊富な音楽の好みを持つ評価者や、音楽経験の少ない評価者を含め様々な音楽経験を持つ評価者に対して定性評価を行い、評価結果を分析する必要があると考えられる。

5. おわりに

本研究では DNN を用いた音楽生成時に、音楽の印象を表す印象ベクトルを入力することで、音楽の印象に沿った和音進行とテンポを生成できるか定性調査によって検討した。入力する印象ベクトルによっては印象を和音進行とテンポに反映できない例も見受けられたが、概ねの印象ベクトルについては入力した印象を和音進行とテンポに反映することができた。今後の課題として、より多様な印象観点の導入や、より豊富な音楽の好みと音楽経験を持つ評価者の確保が挙げられる。

参考文献

[1] Allan, M. and Williams, C.: Harmonising chorales by probabilistic inference, *Advances in neural information processing systems*, pp. 25–32 (2005).
 [2] Raphael, C. and Stoddard, J.: Harmonic analysis with probabilistic graphical models, *ISMIR* (2003).
 [3] Lim, H., Rhyu, S. and Lee, K.: Chord Generation from Symbolic Melody Using BLSTM Networks, *In Proc. of ISMIR*, pp. 621–627 (2017).
 [4] Makris, D., Kaliakatsos-Papakostas, M., Karydis, I. and Ker-manidis, K. L.: Combining LSTM and feed forward neural networks for conditional rhythm composition, *In Proc. of*

EANN, pp. 570–582 (2017).

[5] Shneiderman, B.: Creating creativity: user interfaces for supporting innovation, *ACM Transactions on Computer-Human Interaction (TOCHI)*, Vol. 7, No. 1, pp. 114–138 (2000).
 [6] 梅村祥之: 規則的に生成した 4 音符からなる楽曲を用いた楽曲の心地よさに関する客観評価指標, *研究報告音楽情報科学 (MUS)*, Vol. 2014, No. 17, pp. 1–7 (2014).
 [7] 倉島 研, 金地美知彦, 畑山俊輝: 楽曲の印象と好みを与えるテンポの影響, *研究報告音楽情報科学 (MUS)*, Vol. 2004, No. 111, pp. 125–130 (2004).
 [8] Hochreiter, S. and Schmidhuber, J.: Long short-term memory, *Neural computation*, Vol. 9, No. 8, pp. 1735–1780 (1997).
 [9] 外崎幹二, 島岡 譲: 和声の原理と実習, 音楽之友社 (1958).
 [10] 岩宮眞一郎, 中嶋としえ: サイン音に和音を用いることの効果の検討, *人間工学*, Vol. 45, No. 6, pp. 329–335 (2009).
 [11] Ritossa, D. A. and Rickard, N. S.: The relative utility of ‘pleasantness’ and ‘liking’ dimensions in predicting the emotions expressed by music, *Psychology of Music*, Vol. 32, No. 1, pp. 5–22 (2004).
 [12] Russell, J. A.: A circumplex model of affect., *Journal of personality and social psychology*, Vol. 39, No. 6, p. 1161 (1980).
 [13] Sutskever, I., Vinyals, O. and Le, Q. V.: Sequence to sequence learning with neural networks, *In Proc. of NIPS*, pp. 3104–3112 (2014).
 [14] Donahue, C., Mao, H. H. and McAuley, J.: The NES Music Database: A multi-instrumental dataset with expressive performance attributes, *In Proc. of ISMIR*, pp. 475–482 (2018).
 [15] Burgoyne, J. A., Wild, J. and Fujinaga, I.: An Expert Ground Truth Set for Audio Chord Recognition and Music Analysis, *In Proc. of ISMIR*, pp. 633–638 (2011).
 [16] Scheffe, H.: An Analysis of Variance for Paired Comparisons, *Journal of the American Statistical Association*, Vol. 47, No. 259, pp. 381–400 (1952).