

# 盛り上がり度に基づくループシーケンサにおけるユーザ適応の試み

岩本 祐輝<sup>1,a)</sup> 北原 鉄朗<sup>1,b)</sup>

概要：本稿では、ユーザの嗜好を学習して、ユーザが気に入しやすい曲を自動で作成する作曲ツールの実現を目指す。「ループシーケンサー」という音楽制作ソフトウェア内の音素材を分析して、似たような傾向を持つ音素材を一纏めにし、ユーザが手動で選択した音素材の属するトピック（話題）内の音素材が自動で作曲する際に選ばれやすくするようにした。実験用に簡略化したプログラムを協力者に使っていただき、協力者の気に入った音素材を選択させたところ、自動作曲を行った時に協力者が手を加えた部分に関しては概ね協力者の気に入った音素材が配置される頻度が高くなったと感じたという回答を得ることが出来た。

キーワード：ループシーケンサー トピックモデル LDA ユーザ適応

## 1. はじめに

昔の時代の作曲家は手作業で楽譜を書いて作曲を行っていたが、時代が進んでDTMやDAW等の作曲ツールの登場でPC上で作曲を行うことが出来るようになった。しかし、依然として知識のない人が作曲をしようとするのは難しいままである。作曲自体には興味があるが、作曲の方法や定石を勉強する時間を確保するのが難しく、作曲に手を出すのを躊躇する人が少なからず存在すると思われる。そのため、専門知識がなくても計算機が作曲してくれるシステムの開発が行われているが、ユーザの嗜好を十分に反映できるものになっているとは限らない。

作曲ツールには様々な音が用意されているが、実際にユーザが使う時は全ての音が満遍なく選択されるというのは中々起こりえないと思われる。ユーザが作曲を進めていくと徐々に選択される頻度の高い音と低い音のグループが出来る。そして、グループ内の音は似たような傾向の音が集まっていると考えられる。作曲ツールにユーザの選びやすい音とそうでない音の傾向を学習をさせることで、自動作曲をする際の選択される音に偏りを生じさせて、ユーザの気に入しやすい曲が作られやすくなると考えられる。以上を踏まえて本研究では、ユーザの嗜好を学習して、ユーザが気に入しやすい曲を自動で作成する作曲ツールの実現

を目指す。

既存関連製品 [1] では、AI がサンプルを繋げて楽曲を作りつつ再生していく形となっていて、ユーザは Good と Bad の 2 つの評価をすることでユーザの気に入しやすいものが生成されやすくなるものとなっている。関連研究 [2] では、楽譜表現の遺伝子表現として木構造を用いた複雑なものとなっている。初期の木構造群を生成するために、基準となる音の他に木の深さ、使用する関数と出現頻度等のパラメータを事前に設定して、気に入った音列を複数生成する。生成したものを元に再び木の深さなどを設定して気に入った音列を複数得るというサイクルになっている。[3] は初期状態のモデルを組み込むためのデータを 300 人に及ぶ車両ユーザへのアンケートの回答から集めている。[4] はインターネットのアクセス履歴からユーザの好みを推定して、好みの似ている他のユーザのアクセス履歴からユーザの好みに合うと思われるものを推薦するものである。[5] はユーザの入力した色に対する感性語とユーザのデータベースから好みを推定する。[6] は状態遷移確率を用いて、前回の検索した曲の属性値を基に次の曲の属性値を確率的に予測する手法を取っている。

これらの研究はユーザの評価がどれだけ反映しているかが分からない、参照と修正が GUI で行えない、事前のデータを集める負担が大きいといった点がある。本研究は事前にループシーケンサー内の音素材を分析して、一纏めにして、それをデータとして扱っている。これにより新しく音素材が追加されても、その分のデータの収集の負担は小さ

<sup>1</sup> 日本大学文理学部

<sup>a)</sup> iwamoto@kthrlab.jp

<sup>b)</sup> kitahara@kthrlab.jp

くなる．また，GUI 上で必要な操作を行えるようにして  
 いて，ユーザの評価がどれだけ反映されているかを GUI 上  
 で確認できるようにもしている．そして，本研究のプログ  
 ラムはユーザ個人の音素材の選び方だけを参照して，他の  
 ユーザを参照したり何かを入力する必要は無い．加えて，  
 音素材を自動で配置する時の音素材の選ばれ方は，状態遷  
 移確率の代わりに構成比率を用いている．

## 2. 提案システムの概要



図 1 提案システムの画面

### 2.1 基本構想

元のループシーケンサーを用いたシステム [7] では自動  
 作曲を行う際に，ユーザが手動で描いた曲線（以下，盛り  
 上がり度曲線と呼称）に基づいて，事前に計算した盛り上  
 がり度という値を付与された音素材が配置される（図 1 参  
 照）．音素材は [Sequence][Synth][Bass] [Drums] の 4 パ  
 ートに分けられていて，盛り上がり度は 0~4 の 5 つに分け  
 られている．音素材は 4 つのパート・5 つの盛り上がり度の  
 合計 20 種類のグループに分類されている．同じパート内  
 の同じ盛り上がり度の音素材は等確率でランダムに選ば  
 れるようになっているが，ランダムで選ばれる状態ではユ  
 ザの好みに合う曲は作られにくい．

ユーザの好みに合う曲を作るためには，ユーザの好みに  
 合う音素材が選ばれるようにする必要があり，そのために  
 は音素材の選ばれ方に偏りを生じさせる必要がある．そこ  
 で，音素材の選ばれ方に偏らせるための基準としてユーザ  
 の嗜好を用いることにした．しかし，音素材の数が増える  
 と全ての音素材にユーザの好みに合うかどうかの情報を人  
 力で付与するのは難しくなる．

本研究に於いては，ループシーケンサーを用いての作曲  
 中にユーザの選択する音素材並びに選択しない音素材の傾  
 向をユーザの嗜好と定義している．ループシーケンサー内  
 の数々の音素材を分析して似たような傾向を持つ音素材を  
 分類することで，ユーザが任意の音素材を選択した時に，  
 それ自体とそれと似たような傾向を持つ音素材が自動作曲  
 をする際に選ばれやすくなる．

これにより，自動作曲をした際にユーザが好んで選ぶ音  
 素材が選出されやすくなり，ユーザの好みに合う曲が生成  
 されやすくなると思われる．

### 2.2 提案手法の事前準備

#### 2.2.1 メルスペクトログラム

音声分類を行う際にはメルスペクトログラムの方が高速  
 フーリエ変換で得られたものを可視化したものより分類の  
 精度が良いという報告があるので [9]，本研究では音素材  
 からメルスペクトログラムという特徴量を抽出して分類を  
 行っていく．メルスペクトログラムは行列の形を取り，行  
 が周波数，列が時間軸，値はパワーを表す．

#### 2.2.2 K-平均法

20 種類に分類されている音素材のグループの内，任意の  
 グループに存在する全ての音素材のメルスペクトログラム  
 の行列を，行数を増やす形で連結させた後に転置した行列  
 を得る．その行列に対してクラスタ数を 8 として K-平均  
 法 [10] を用いてクラスタリングを行い分類結果が格納され  
 た配列を得る．得た配列を等分割することで，そのグル  
 ープ内の各音素材の各時間軸毎の特徴量がどのクラスタに  
 属しているかの情報を得ることが出来る．どのクラスタに何  
 個属しているかを数え上げたものを，後述のトピックモデ  
 ルに適用するための文書データの形式で纏める．

#### 2.2.3 トピックモデル

文書データを解析する手法としてトピックモデルという  
 ものがある [11]．データの集合の背景に話題等（トピック）  
 が存在して，データはそれに従って分布される．その上で  
 1 つのデータが複数のトピックを持っている．そう仮定し  
 たものに対して，そのデータがそれぞれのトピックの中か  
 らあるトピックに属している確率を求める手法である．本  
 研究では，その手法の一つの LDA [12] に対して前述で得た  
 データを用いて分析をする．今回は 4 つのトピックで分類  
 を行った．

	topic0	topic1	topic2	topic3
Knarzer Beat D	0.001685633351	0.9948140041	0.001685633351	0.001814729157
Hard Beat A	0.001685838454	0.001823552217	0.001685838454	0.9948047709
Knarzer Beat A	0.001685669733	0.9948160649	0.001685669733	0.001812595672
Knarzer Beat H	0.00168564623	0.994815314	0.00168564623	0.001813393572
Knarzer Beat B	0.001685595582	0.9948171144	0.001685595582	0.001811694392
Knarzer Beat F	0.001685633351	0.9948140041	0.001685633351	0.001814729157
Knarzer Beat J	0.00168564623	0.994815314	0.00168564623	0.001813393572

図 2 音素材のトピック比率のデータの一例（行：音素材の名前，列：  
 トピック）

### 2.3 提案手法

トピック数を 4 として，20 種類に分類されている音素  
 材のグループの内，任意のグループに属する音素材の各ト  
 ピックに対する選好度を  $\alpha = (\alpha_1, \alpha_2, \alpha_3, \alpha_4)$  とし  
 て，初期

値を一律 1.0 とする．ある音素材  $x$  が手動で配置された時， $x$  のトピック比率を  $z = (z_{x1}, z_{x2}, z_{x3}, z_{x4})$  とすると，各トピックに対する選好度は以下の式のように更新される．

$$\alpha' = z + \alpha$$

あるグループ内の音素材が  $l$  個存在する時，任意の音素材  $x$  が選ばれる確率は，以下の式になる．

$$\frac{\sum_{n=1}^4 \alpha_n z_{xn}}{\sum_{k=1}^l (\sum_{n=1}^4 \alpha_n z_{kn})}$$

例を挙げると，あるグループ内での各トピックに対する選好度  $\alpha$  の内，1 つ目の値が大きければ，トピック比率  $z$  の 1 つ目の値が大きい音素材が選ばれやすくなるようになる．

## 2.4 プログラムの説明

### 2.4.1 画面の説明

本研究室では以前より，ユーザが盛り上がり度の曲線を描くと，それに合わせて音素材を自動で挿入するループシーケンサーの開発を行ってきた．プログラム全体は図 1 の様になっている．

このシステムでは，図 1 の中央部分でユーザが盛り上がり度曲線を手動で描くと，一定間隔毎に平均の盛り上がり度を計算して，[Sequence][Synth][Bass][Drums] の 4 パートのそれぞれに音素材が入るかどうかを決めた後に，事前に各音素材に自動で計算して割り振られた盛り上がり度と平均の盛り上がり度を比較して，図 1 の中央下部分に自動で音素材を割り振られるように出来ている．

図 1 の左上の部分の [control] 枠内の [create] ボタンを押すとユーザの描いた盛り上がり度曲線に基づいて図 1 の中央下部分に音素材が配置される．図 1 の中央下部分の任意の長方形を左クリックすると黒い枠で囲まれて，その部分が選択されていることが可視化される．そして，選択箇所に対応したパートに応じて図 1 の左下側部分に，対応したパート内の音素材が格納された円が盛り上がり度に応じた色分けで表示される．任意の円を一度左クリックすることで対応した音素材の音声再生が再生されて，もう一度同じ円を左クリックすることで図 1 の中央下部分で選択した部分に，対応した音素材が挿入される．

図 3 と図 4 は図 1 の左中央部分を拡大したもので，節 2.3 で述べた選択された音素材のトピック比率（左側）と選択された音素材が属するグループの各トピックに対する選好度（右側）を可視化したものである．ユーザが手動で音素材を配置することで，配置された音素材のトピック比率に従ってグループ内の各トピックに対する選好度が変動して，自動作曲の際の音素材の選ばれ方に偏りが生じるようになる．

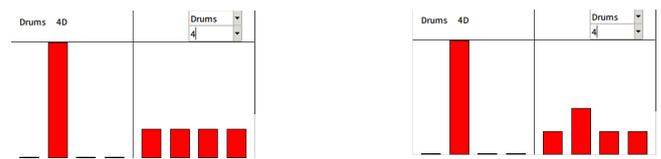


図 3 音素材挿入前のトピック比率 図 4 音素材挿入後のトピック比率とトピックに対する選好度

### 2.4.2 一連の流れ

図 5 は手順 3 を実行時の画面である．

- (1) 盛り上がり度曲線を任意に描き，[create] ボタンをクリックして自動で作曲をする．
- (2) 任意の配置された音素材を左クリックして選択する．
- (3) 画面左下に表示される各音素材に対応した丸を左クリックして音素材を入れ替える．
- (4) ある程度入れ替えたら，任意で盛り上がり度曲線を書き換えて，[create] ボタンをクリックする．
- (5) 2 から 4 までを繰り返す．

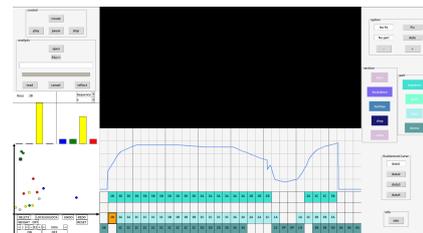


図 5 手順 3 部分

## 3. 評価実験

### 3.1 実験

実験は 2 回に分けられて，1 回目は匿名の 7 人に協力していただき，2 回目は匿名の 2 人に協力していただいた．

#### 3.1.1 実験の設定

今回の実験では，図 1 の右上にある [option] 枠内のボタンの設定を変えている．

音素材を配置する部分は 8 小節分あり，1 小節内で各パート毎に 4 つの音素材を配置することが出来る．[No fix] ボタンが押されている時は，1 小節内の各パートの各ブロック毎に音素材を配置するかどうかを決めるが，[Fix] ボタンが押されている時は，1 小節内の各パート毎に音素材を 1 つも配置しない，もしくは 4 ブロックとも同じ音素材で配置するかを決めるようになる．[No part] ボタンが押されている時は，自動作曲をする際に盛り上がり度曲線に基づいて音素材を配置するだけだが，[Auto] ボタンが押されている時は，それに加えて曲の構成を考慮に入れて自動作曲をするようになる．

今回の実験では [Fix] ボタンと [Auto] ボタンが押されている状態に設定した．それに加えて手動で音素材を配置した時は，一括で 1 小節内の 4 ブロックに選択した音素材を

配置して、[LOCK\UNLOCK] ボタンをクリックした時は、一括で選択した小節内の選択したパート部分の4ブロックのロック状態のON\OFFの切り替えを出来るようにした。

### 3.1.2 実験の流れ

節4の設定で実験を行うことにした。実験は2回に分けられて、2つの盛り上がり度曲線をそれぞれの実験で用いる。実験の協力者には、指定した以下の手順でプログラムを操作するように指示をした。また、実験2回目のトピック選好度は、実験1回目終了時のトピック選好度をそのまま引き継いで行うものとした。また図6と図7の矢印は、その矢印が指している[Sequence]パート、または[Drums]パートの配置されている音素材を置き換えることを意味している。

- (1) 1回目の実験用または2回目の実験用の盛り上がり度曲線が描かれているのを確認してから自動作曲を実行して、生成された曲を試聴して評価を行う。
- (2) [Sequence]パートの音素材を協力者の気に入った音素材に置き換えた後、もう一度曲を試聴して評価を行う。
- (3) 自動作曲を実行後、生成された曲を試聴して評価を行う。
- (4) [Drums]パートの音素材を協力者の気に入った音素材に置き換えた後、もう一度曲を試聴して評価を行う。
- (5) 自動作曲を実行後、生成された曲を試聴して評価を行う。
- (6) 手順1～手順5を最低一回繰り返す。

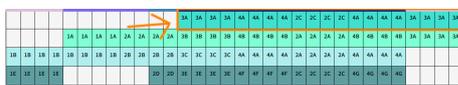


図6 実験手順2部分



図7 実験手順4部分

### 3.1.3 質問の内容

1回目の実験と2回目の実験が終わる度に共通の質問に答える。以下の2つが質問の内容である。

- (1) 実験序盤の自動作曲と実験終盤の自動作曲を比較した際に、ユーザが手で選んだ音素材と選んだものと同じ種類の音素材が選ばれる頻度が高くなったと感じましたか。
- (2) 実験序盤の自動作曲と実験終盤の自動作曲を比較した際に、ユーザの好みに合う曲が生成されやすくなったと感じましたか。

質問1の意図は、提案手法によって自動作曲の際にユーザが選んだ音素材と同じ種類の音素材が選ばれやすくなったかどうかを調べるためである。質問2の意図は、質問1

の回答を基に、実際にユーザの好みに合う曲が生成されやすくなったかどうかを調べるためである。

それぞれの質問に対して感じたかどうかを1~5の5段階で評価をしてもらう。1に近いほど感じず、5に近いほど感じたことを意味する。また、そのように回答した理由も併せて回答してもらった。

### 3.2 実験結果

1回目の実験に参加して頂いた協力者7人をA, B, C, D, E, F, Gと呼称する。2回目の実験には1回目の実験の協力者A, Bが引き続き参加して頂いた。1回目の実験もしくは2回目の実験が終わる度に節3.1.3の質問に回答してもらった。

#### 3.2.1 1回目の実験結果と考察

以下の表1が協力者7人の回答を纏めたものである。

質問\回答者	A	B	C	D	E	F	G
1つ目	4	4	4	5	4	5	4
2つ目	3	4	3	5	4	2	5

表1 1回目の実験の質問の回答

1つ目の質問の回答を見ると、全員自動作曲の際にユーザが選んだ音素材と同じ種類の音素材が選ばれやすくなったと感じていることが分かった。2つ目の質問を回答を見ると、実際にユーザの好みに合う曲が生成されたと感じた人と感じなかった人で別れる結果になった。

次に、自動作曲で生成された曲もしくは[Sequence][Drums]部分の音素材を入れ替えて出来た曲を試聴した後、試聴した曲のユーザ毎の評価を集計しデータにした。そのデータから一番最初の評価と一番最後の評価のみを抜粋して以下の表2に纏めた。一番最初の評価はトピック選好度が初期値のままの状態で行った自動作曲を行い生成された曲を評価したもので、一番最後の評価は実験終盤でユーザ毎にトピック選好度が変化しきった状態で自動作曲を行い生成された曲を評価したものである。更に手で音素材を配置した回数を以下の表3にユーザ毎に纏めた。

	最初	最後
A	4	4
B	3	3
C	1	4
D	4	2
E	4	4
F	5	4
G	2	5

表2 評価の比較

協力者	挿入回数
A	51
B	34
C	21
D	73
E	51
F	28
G	38

表3 音素材の挿入回数

1回目の実験において、実験手順を最小限に行った場合は音素材の挿入回数は20回になる。2つの表を見比べて見るとC, D, Gが特徴的なものが見て取れる。C, Gの音素材の挿入回数は比較的少なく、一番最初の評価はそれぞれ1, 2と低い。一番最後の評価はそれぞれ4, 5と大幅に評価が上がっている結果となっている。表2のC, Gとそ

れ以外の一番最初と一番最後の評価を見比べてみた結果、トピック選好度が初期値の状態で作られた曲がユーザの好みに合わなかった時は、今回の提案手法が効果を発揮する可能性が高いと推測することが出来る。

反対に D は挿入回数が 73 と群を抜いて多く、評価も 4 から 2 へと大きく下がっている。C、G と D で対称的な結果になった理由を探るべく、以下のそれぞれのトピック選好度の遷移を折れ線グラフで可視化したものを見比べる。折れ線グラフの折れ線は 4 本存在する。色は(青, 赤, 黄, 緑)となっていて、それぞれ(トピック 0, トピック 1, トピック 2, トピック 3)に対応している。

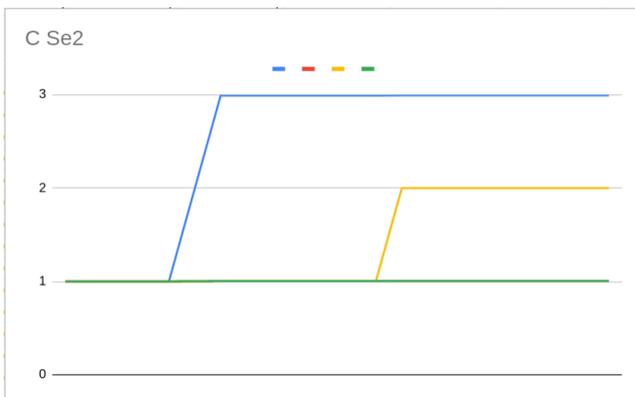


図 8 C の [Sequence] 盛り上がり度 2 のトピック選好度の遷移

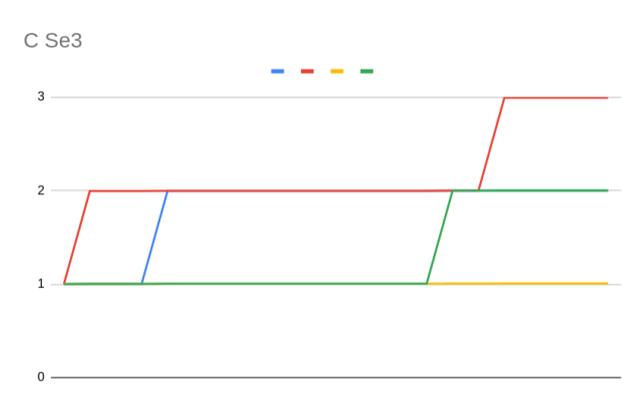


図 9 C の [Sequence] 盛り上がり度 3 のトピック選好度の遷移



図 10 G の [Sequence] 盛り上がり度 2 のトピック選好度の遷移



図 11 G の [Sequence] 盛り上がり度 3 のトピック選好度の遷移

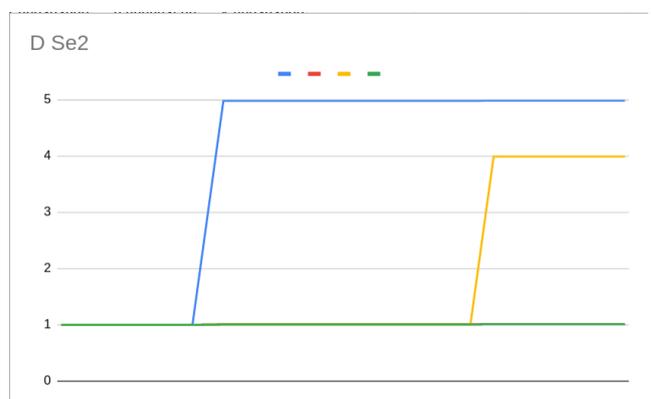


図 12 D の [Sequence] 盛り上がり度 2 のトピック選好度の遷移



図 13 D の [Sequence] 盛り上がり度  
3 のトピック選好度の遷移

実際に見比べて見たところ、一番最初に低い評価を出してから一番最後に高い評価を出したという共通点を持つ C, G の間には一貫した共通点は無く、C, G とは対称的な D は C もしくは G と似たような傾向をしている部分があることが分かった。

例を挙げると、図 8 と図 10, 図 9 と図 11 を見れば、値が増加しているトピックの数や増加の傾向が違うことが明確に分かる。図 8 と図 12 を見れば C と D で同じトピックの値が似たような傾向で増加しているのが分かり、図 11 と図 13 を見れば、G と D でトピックの種類こそ違いますが似たような傾向でトピックの値が増加しているのが分かる。

このことから、C, G の二人の生成された曲に対する評価は似た傾向を示していたが、それぞれの音素材の選び方は違うものだということが分かった。

2 回目の質問で 2 と 3 の評価をした A, C, F の回答理由は「好きな音が入ったものもあったが、あまり好みではない音も同時に入ってしまった」「自分が操作していた [Sequence] と [Drums] は好みに寄って来たが、それに伴って [Synth] と [Bass] が好みに寄って来るとは感じなかった」「選んだ音素材がパートごとに良く感じても、全体として聞いたときに合わないことがあったため」というものだった。

今回の実験では [Sequence][Drums] 部分の音素材だけを入れ替えたので、[Synth] [Bass] 部分の音素材に選ばれ方に変化は生じないようにしている。そのため、自動作曲をした際に [Sequence][Drums] 部分ではユーザの好みに合う音素材と同じトピックに属する音素材が選ばれても、[Synth][Bass] 部分ではユーザの好みに合う音素材が選ばれてしまって、4 パートの音素材が噛み合わなくなり、ユーザの好みに合う曲が生成されたと感じなくなってしまう場合も考えられる。

1 回目の実験は、トピック選好度が初期値の時に生成された曲に対する評価が低い時は今回の提案手法が一定の効果を発揮すると推測することが出来るという結果になった。

### 3.2.2 2 回目の実験

以下の表 4 が A, B の回答を纏めたものである。

質問\回答者	A	B
1 回目	3	4
2 回目	4	4

表 4 2 回目の実験の質問の回答

2 回目の実験は、1 回目の実験終了時のトピック選好度を引き継いで行っており、既に音素材の選ばれ方の偏り具合はある程度収束している形になっているので、1 回目の実験より自動作曲の際にユーザの選んだ音素材と同じトピックに属する音素材が選ばれる頻度が高くなったと感じにくくなると思われる。

しかし、1 回目の実験とは異なる盛り上がり度曲線を用いて自動作曲を行ったことで、1 度も配置されることの無かった [Sequence] 部分の盛り上がり度 1 の音素材が配置されるようになった。それによりユーザは新しくその部分の音素材に手を加えられるようになり、1 回目の実験では変化しなかった [Sequence] 部分の盛り上がり度 1 のトピック選好度に変化を起すことが出来た。その結果、自動作曲の際によりユーザの好みに合う曲が作られる頻度が高くなったと感じるようになったと推測することが出来る。

1 回目の実験と 2 回目の実験を通して、違う盛り上がり度曲線を用いても今回の提案手法は一定の効果を発揮すると推測することが出来るという結果になった。

### 3.2.3 トピック選好度の比較

最後に実験 1 回目が終わった時点での協力者 7 人のトピック選好度を表にしたものを以下に載せる。

パート+盛り上がり度 トピック	0	1	2	3
Sequence0	1.000	1.000	1.000	1.000
Sequence1	1.000	1.000	1.000	1.000
Sequence2	2.005	1.012	6.971	1.012
Sequence3	3.012	12.944	1.025	2.019
Sequence4	2.998	1.012	5.978	1.012
Drums0	1.000	1.000	1.000	1.000
Drums1	5.973	1.009	1.008	1.009
Drums2	1.020	2.527	1.020	11.433
Drums3	3.990	1.010	3.990	1.010
Drums4	1.013	2.008	1.0135	7.965

表 5 A のトピック選好度

パート+盛り上がり度 トピック	0	1	2	3
Sequence0	1.000	1.000	1.000	1.000
Sequence1	1.000	1.000	1.000	1.000
Sequence2	2.991	1.005	1.998	1.005
Sequence3	1.010	3.990	1.010	3.990
Sequence4	4.984	1.012	3.992	1.012
Drums0	1.000	1.000	1.000	1.000
Drums1	1.005	2.991	1.005	1.999
Drums2	1.010	1.011	1.010	6.969
Drums3	1.998	1.005	2.991	1.005
Drums4	1.010	4.983	1.010	2.997

表 6 B のトピック選好度

パート+盛り上がり度	トピック	0	1	2	3
Sequence0		1.000	1.000	1.000	1.000
Sequence1		1.000	1.000	1.000	1.000
Sequence2		2.991	1.005	1.998	1.005
Sequence3		2.000	2.993	1.007	2.000
Sequence4		2.995	1.008	3.988	1.008
Drums0		1.000	1.000	1.000	1.000
Drums1		1.995	1.002	1.002	1.002
Drums2		1.003	1.004	1.003	2.990
Drums3		1.004	1.003	2.990	1.003
Drums4		1.007	4.979	1.007	1.007

表 7 C のトピック選好度

パート+盛り上がり度	トピック	0	1	2	3
Sequence0		1.000	1.000	1.000	1.000
Sequence1		1.000	1.000	1.000	1.000
Sequence2		4.985	1.012	3.992	1.012
Sequence3		13.934	2.018	1.024	1.024
Sequence4		4.999	1.027	12.947	1.027
Drums0		1.000	1.000	1.000	1.000
Drums1		1.005	1.005	1.005	3.985
Drums2		1.010	4.774	1.010	3.206
Drums3		7.967	1.015	3.002	1.015
Drums4		1.030	11.956	1.030	7.984

表 8 D のトピック選好度

パート+盛り上がり度	トピック	0	1	2	3
Sequence0		1.000	1.000	1.000	1.000
Sequence1		1.000	1.000	1.000	1.000
Sequence2		6.971	1.012	2.006	1.012
Sequence3		3.002	5.982	1.015	3.002
Sequence4		2.007	1.013	7.966	1.013
Drums0		1.000	1.000	1.000	1.000
Drums1		4.981	1.009	1.008	2.002
Drums2		1.017	3.276	1.017	8.690
Drums3		2.993	1.007	2.993	1.007
Drums4		1.012	2.006	1.012	6.971

表 9 E のトピック選好度

パート+盛り上がり度	トピック	0	1	2	3
Sequence0		1.000	1.000	1.000	1.000
Sequence1		1.000	1.000	1.000	1.000
Sequence2		2.990	1.003	1.004	1.003
Sequence3		3.988	2.995	1.008	1.009
Sequence4		2.993	1.007	2.994	1.007
Drums0		1.000	1.000	1.000	1.000
Drums1		1.998	1.005	1.005	2.992
Drums2		1.007	1.760	1.007	4.227
Drums3		3.984	1.005	1.005	1.005
Drums4		1.012	5.978	1.012	2.999

表 10 F のトピック選好度

パート+盛り上がり度	トピック	0	1	2	3
Sequence0		1.000	1.000	1.000	1.000
Sequence1		1.000	1.000	1.000	1.000
Sequence2		1.005	1.005	3.985	1.005
Sequence3		1.014	2.007	1.013	7.965
Sequence4		1.010	1.010	6.970	1.010
Drums0		1.000	1.000	1.000	1.000
Drums1		3.624	1.007	1.007	2.362
Drums2		1.003	2.509	1.003	1.485
Drums3		2.990	1.003	1.004	1.003
Drums4		1.022	4.996	1.022	9.960

表 11 G のトピック選好度

4つのパートと5つの盛り上がり度の合計20グループ毎にトピック選好度が設けられているが、グループ毎に音素材を4つのトピックに分類しているため、同じトピックでもグループ毎の音素材の傾向は変わってくる。聴き比べてみると、盛り上がり度2の[Drums]のトピック3と盛り上がり度3の[Drums]のトピック2と盛り上がり度4の[Drums]のトピック1に分類されている音素材の傾向は似ているのが分かった。

それを踏まえて、表5,6,7,8,9,10,11のA,B,C,D,E,F,Gの[Drums]部分のトピック選好度を見てみると、例外こそあるが何れかの1つが高い数値を取ると他の2つの内のどちらか、もしくは両方共高い数値を取っていることが分かる。少なくとも、1回目の実験に協力してくれた人達は似たような音素材を進んで選んでいることが分かった。

#### 4. おわりに

本稿では提案手法として、事前にループシーケンサー内の4つのパート内の5つの盛り上がり度に分類された音素材を分析して、4つのトピックという属性に分類を行った。ユーザーが手で音素材を選んで配置する度に、その音素材が属するトピックが選ばれやすくなるようにすることで、ユーザーが選んだ音素材と似た傾向の未参照の音素材も選ばれやすくなり、また音素材の数が増えたときにも対応できるようになった。

本研究の実験用に簡略化したプログラムを用いた実験を

2回に分けて実施を行ったところ, 1回目の実験には7人が, 2回目の実験には2人が協力してくれた。1回目の実験では, 各音素材の選ばれる確率が等しい状態で行われた自動作曲により作られた曲が気に入らなかった場合は, 提案手法を用いることで最終的にユーザの好みに合う曲が作られやすくなるという推測が出来る結果になった。また, 1回目の実験と2回目の実験を通して, 盛り上がり度曲線を変えた状態でも提案手法は一定の効果を示す結果になった。

今回の実験では [Sequence][Drums] 部分の音素材のみに手を加えていたので, [Synth][Bass] 部分にも手を加える実験と, 音素材の数を増やした状態でも今回の提案手法が通じるかどうかの実験を行いたい。また, ユーザが手動で音素材を挿入した回数等のデータを収集する方法の精度が悪かったので, より色々なデータを細かく収集できるようにしていきたい。

本システムが, 専門知識の無い人でも簡単に好きな曲を作れるようになる画期的な作曲システムを実現するための一助になることを期待したい。

謝辞 本研究は, 科研費 19K12288, 16H01744, 21H03572 の支援を受けて行われた。

#### 参考文献

- [1] Mubert - Thousands of Staff-Picked Royalty-Free Music Tracks for Streaming, Videos, Podcasts, Commercial Use and Online Content, 入手先 (<https://mubert.com/>), (参照 1/10 21:09)
- [2] 安藤 大地, Dahlstedt Palle, Nordahl Mats, 伊庭 斉志: 対話型 GP を用いたクラシック音楽のための作曲支援システム, 芸術科学会論文誌, Vol.4, No.2, pp.77-86, 2005.
- [3] 岩崎 弘利, 水野 伸洋, 原 孝介, 本村 陽一: ユーザの好みに合わせてコンテンツを推薦するカーナビへのページアンネットの適用, 電子情報通信学会技術研究報告, 104(226), pp.25-30. 2004.
- [4] 井原 雅行, 金田 洋二, 上野 圭一, 金山 英明: ユーザの潜在的好み推定法, 電子情報通信学会論文誌 A, Vol.J82-A, No. 5, pp.717-725, 1995.
- [5] 柳生 智彦, 久森 芳彦, 八木 康史, 谷内田 正彦: 配色支援システムにおける好みの獲得と迷いの解消, 電子情報通信学会論文誌 A, Vol.J79-A, No.2, pp.261-270, 1996.
- [6] 井原 雅行, 金田 洋二, 上野 圭一, 金山 英明: 情報をもつ属性を用いたユーザ動的好み推定法の提案およびその評価, 電子情報通信学会論文誌 A, Vol.J80-A, No. 4, pp.700-707, 1997.
- [7] 飯島孔右, 鶴岡亜矢佳: 手書き入力で盛り上がりをコントロールするループシーケンサ: スペクトログラムから盛り上がりの自動割り振り, 第 77 回全国大会講演論文集, Vol.2015, No.1, pp.373-374, 2015.
- [8] Cooley, James W., Tukey, John W.: An algorithm for the machine calculation of complex Fourier series, Math.Comput., 19 (90), pp.297-301, 1965.
- [9] Bonaventure F. P. Dossou, Yeno K. S. Gbenou :FSER: Deep Convolutional Neural Networks for Speech Emotion Recognition, Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV) Workshops, pp.3533-3538, 2021.
- [10] MacQueen, J. B.: Some Methods for classification and

Analysis of Multivariate Observations, Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability. University of California Press, pp.281-297, 1967.

- [11] Thomas Hofmann: Probabilistic Latent Semantic Indexing, International Computer Science Institute, Berkeley, CA & EECS Department, CS Division, UC Berkeley, 1999.
- [12] Blei, D. M., Ng, A.Y. and Jordan, M.I.: Latent Dirichlet Allocation. Journal of Machine Learning Research, Volume 3, pp.993-1022, 2003.