

楽譜コンテキストのベイジアンマイニングに基づく 自動伴奏付与システム

- ユーザの感性的要求に応じた伴奏生成 -

多田 圭吾^{†1} 岡村 亮吾^{†1}
山西 良典^{†1} 加藤 昇平^{†1}

近年, Desktop Music (DTM) 関連機器やソフトウェアなどの発展・普及により, 音楽を嗜む一環として作曲が注目されているが, 作曲は高度な知識や経験を必要とする. そこで本研究では, 楽譜コンテキストのベイジアンマイニングに基づく自動伴奏付与システムを提案する. 提案システムでは, 音楽の時間的変動を楽譜コンテキストとして扱い, ベイジアンマイニングを用いて学習曲における音楽構成要素間の因果関係を学習する. 提案システムは, ユーザの音楽経験の有無に関わらず, 任意のニュアンスが付与された伴奏を生成する. 伴奏生成実験および主観評価実験によって, 入力されたメロディに応じて学習曲のジャンルのニュアンス, および, 印象的ニュアンスが付与された伴奏が生成されることを確認し, また既存システムとの比較実験により提案システムの高い有用性を確認した.

Accompanying System Based on Bayesian Mining of Musical Score Context: Ondemand Accompany for Affective Requests

KEIGO TADA,^{†1} RYOGO OKAMURA,^{†1}
RYOSUKE YAMANISHI^{†1} and SHOHEI KATO^{†1}

Present days, more public interested in music focuses on composing as one of the forms to enjoy music. However, it is difficult for people who does not have sufficient musical experiences. Thus, we proposed an automated accompanying system focusing on bass and drums by using Bayesian mining of the score context. In this study, we explained the musical temporal variates as the score context, and learned the nuance of the music in the database using Bayesian mining. In a composing experiment, we obtained various accompanies depending on the music database used for learning, and the nuance of the music in the

database. And, we confirmed the availability of the proposed system through three types of subjective evaluation experiment. We believe that this system enable us to compose music with accompany that has the nuance of the music in the database, even if user does not have much musical experiences or not.

1. はじめに

今日, 生活のあらゆる場面に多種多様なエンターテイメントが存在している. 特に音楽は古くから世界中の人々に愛され, 聴取, 演奏, 制作など様々な形で楽しまれてきたメディアの一つである. 情報科学分野においても機械学習を用いた音楽認識 [1-4] や音楽生成システム [5, 6] などが注目されている. また様々な音楽作成ソフトウェアが販売されており [7-9], 音楽制作は多くの人にとってより身近なものとなったといえる.

伴奏は音楽の印象を決定する上で非常に大きな役割を担う [10] が, 伴奏作成に必要な複数の音楽的要素の相互作用を考慮する為には高度な知識と多大な経験が必要であり, 音楽経験が不十分な人々には最大の難点の一つといえる. そこで, 本研究では音楽的知識や経験を必要としない作曲支援システムの実現を目指し, 自動伴奏付与システムを提案する [11]. 提案システムでは用意したデータベース内にある楽曲の構成, すなわち, 複数の楽器の相互関係をベイジアンネットワーク [12] を用いて学習する. 学習した因果関係を基に提案システムは入力されたメロディに, 学習対象となった楽曲のニュアンスをもつ伴奏を付与する.

本稿では, 提案システムが学習曲のニュアンスをもつ伴奏が生成可能であることを確認するため, 音楽のジャンルごとに楽曲を学習し生成された伴奏の比較を行なった. また, 同様の実験を「明るい」や「激しい」などの人間が楽曲から感じる印象についても行うことで, 提案システムが音楽のジャンルのみでなく, 楽曲の印象的ニュアンスを学習しそれに応じた伴奏を付与出来ることを確認する.

2. 提案システムの概要

図 1 に提案システムの概要を示す. 本システムでは予め楽曲を構成する複数の楽器の相互関係をベイジアンネットワークを用いて学習したモデルに, 伴奏を付与したい楽曲のメロディ, コード, およびセクションの情報を入力することで, 入力したメロディに応じたベ-

^{†1} 名古屋工業大学大学院工学研究科情報工学専攻

Dept. of Computer Science and Engineering, Graduate School of Engineering, Nagoya Institute of Technology University

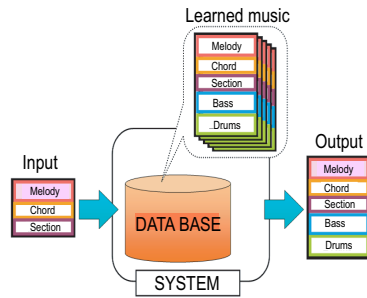


図 1 提案システムの概要

ストとドラムが伴奏として付与される．本稿では，音楽の持つニュアンスに大きく影響を与える楽器としてベースとドラムに着目している．

本研究では，楽曲の楽譜情報における時間的前後関係を表現した楽譜コンテキストから，複数の楽器のメロディやリズム，音高などの音楽の構成要素間の因果関係を学習する．ユーザは作曲に関する知識や経験を必要とせず，既存曲を学習曲とすることで，学習に用いた楽曲のニュアンスが付与された伴奏を生成できる．

3. 楽譜コンテキスト

図 2 に本稿で定義する楽曲の構造を示す．セクションは楽曲を構成する小節の集合であり，それぞれの小節が持つメロディとコードの類似性により決定・分類される [13, 14]．セクションはブリッジやモチーフ，コーラスとして広く認識されており，作曲における重要な概念であると言われている．しかしながら，セクションについては一般的な定義が無いため，本稿ではセクションはヒューリスティックに決定・分類した．

学習には，4 分の 4 拍子であり，メロディとコード，ベース，ドラムから構成される楽曲を用いた．また，メロディやベースなどの音符の最小単位は 16 分音符とし，これを 1 拍とする．

本稿では，楽譜コンテキストはコード，メロディのリズム，メロディの音高，ベースのリズム，ベースの音高，ドラム，小節カウンタの 7 要素により構成される．小節 t の p 拍目における楽譜コンテキスト sc_p^t は，下式で表される．

$$sc_p^t = (ch_p^t, mr_p^t, mp_p^t, br_p^t, bp_p^t, dr_p^t, bc_p^t) \quad (1)$$

楽譜コンテキストを構成する各要素については下節で詳細を示す．

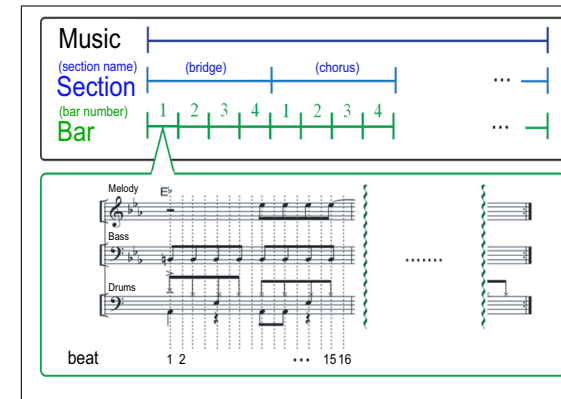


図 2 本稿における楽曲の構造

3.1 コードのモデル化

コードは楽曲の任意の時刻における和音構成であり，メロディやベースがとりうる音高は基本的にコードの構成音に准ずるものとなる．本稿では，小節 t の p 拍目におけるコード状態を ch_p^t とし，コード名を用いて表す．コード名にはコードルートと調性を表す記号を用いた [15]．

3.2 メロディとベースのモデル化

メロディおよびベースは，リズムと音高に分けてそれぞれモデル化する [16]．

3.2.1 リズムのモデル化

小節 t の p 拍目におけるメロディとベースのリズム状態をそれぞれ mr_p^t, br_p^t とする．リズムは発音を表す音符と無音を表す休符によって構成されており，RS1：発音，RS2：休符，RS3：前状態を維持，の 3 状態によって表現される．

3.2.2 音高のモデル化

小節 t の p 拍目におけるメロディとベースの音高をそれぞれ mp_p^t, bp_p^t とし，コードに対する相対音高を用いて表現する．ここで，コードに対する相対音高とは小節 t の p 拍における音高と，小節 t におけるコードの基準音高との度数を示す（表 1 参照）．メロディとベースがとりうる音高は，Musical Instrument Digital Interface (MIDI) の音高情報に基づき 0 から 127 の数字で表現する．また，リズム状態が RS2 であるとき，音高は ϕ とする．

表 1 各コードの基準音

コード	C	C #	D	D #	E	F	F #	G	G #	A	A #	B
コードの基準音	60	61	62	63	64	65	66	67	68	69	70	71

CRASH CYMBAL	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
RIDE CYMBAL	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
OPEN HIHAT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
CLOSED HIHAT	0	0	1	0	1	0	1	0	1	0	1	0	1	0	0	
TOM	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
FLOOR TOM	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
SNARE	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	
BASS DRUM	1	0	0	0	0	0	0	0	1	0	1	0	0	0	0	
p	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
dr_p^t	81	00	10	00	12	00	10	00	11	00	11	00	12	00	10	00

図 3 ドラムパターン例とそのドラム状態 dr_p^t

3.3 ドラムのモデル化

小節 t の p 拍目におけるドラム状態を dr_p^t とし、ドラムを構成する打楽器の発音状態の組み合わせによって表現する。本稿ではドラムを構成する楽器としてクラッシュシンバル、ライドシンバル、オープンハイハット、クローズドハイハット、タム、フロアタム、スネア、バスドラムの 8 つの打楽器を用いる。打楽器には音長の概念がないため、ドラムの状態は RS1 と RS2 (3.2.1 節参照) のみで表現する。図 3 にドラムパターンの例と対応したドラム状態 dr_p^t を示す。

3.4 小節カウンタ

1 セクション内における小節番号を小節カウンタとし、小節 t の p 拍目における小節カウンタを bc_p^t とする。セクションが次のセクションへ推移するとき、小節カウンタはリセットされる。 bc_p^t はその定義上、同小節内の全ての拍で同じ値を持つこととなる。

楽譜コンテキストは前述したコード、メロディのリズム、メロディの音高、ベースのリズム、ベースの音高、ドラム、そして、小節カウンタの 7 要素で表現される。但し、小節 t に

p	1	2	3	4	5	6	7	8	9	...
ch_p^t	Dmaj	Dmaj	Dmaj	Dmaj	Dmaj	Dmaj	Dmaj	Dmaj	Dmaj	...
mr_p^t	RS1	RS3	RS3	RS3	RS1	RS3	RS3	RS3	RS2	...
mp_p^t	11	11	11	11	11	11	11	11	ϕ	...
br_p^t	RS1	RS2	RS1	RS2	RS1	RS2	RS1	RS2	RS1	...
bp_p^t	-10	-10	-10	-10	-10	-10	-10	-10	-10	...
dr_p^t	81	00	10	00	12	00	10	02	10	...
bc_p^t	1	1	1	1	1	1	1	1	1	...

図 4 楽譜コンテキストの一例

おけるドラムまたはベースに全休符が含まれるとき、その楽譜コンテキストは学習対象外とする。図 4 に小節 t における楽譜コンテキスト例を示す。

4. 伴奏生成モデル

本システムでは前節で示した楽譜コンテキストから算出されたパラメータを用いて、入力メロディに対するベースのリズム、ベースのメロディ、ドラムのそれぞれの生成モデルをベイジアンネットワークによって構築する。

4.1 ベースのリズム生成

学習曲におけるベースのリズムと以下の 3 つのパラメータとの因果関係を学習し、ベースのリズム生成モデルを構築する。

mr : メロディのリズム状態

$partiton$: 小節内での拍番号

bc : 小節カウンタ

4.2 ベースの音高生成

本研究では、生成したベースのリズム状態が RS1 であるとき音高を生成する。学習曲に

おけるベースの音高と以下の7つのパラメータとの因果関係を学習し、ベースの音高生成モデルを構築する。

length : 音長

cadence : 楽曲の主調音に対するコードルートの相対度数

pre_cadence : 一つ前の拍に関する *cadence*

next_cadence : 次の拍に関する *cadence*

chord_key : コードの長調・短調

partition : 小節内での拍番号

bc : 小節カウンタ

4.3 ドラムの生成

学習曲におけるドラムと以下の4つのパラメータとの因果関係を学習し、ドラムの生成モデルを構築する。

mr : メロディのリズム状態

br : ベースのリズム状態

partition : 小節内の拍番号

bc : 小節カウンタ

5. 伴奏生成実験

本稿では、2種類の基準（ジャンル、印象）によって構築した伴奏生成モデルを用いた伴奏生成および主観評価実験により、提案システムを用いることで学習曲のニュアンスをもつ伴奏の生成が可能であるか確認する。また、入力メロディに伴奏を付与した出力楽曲の主観評価について既存の伴奏生成システムと比較実験を行い、提案システムの有効性を確認する。

5.1 ジャンル学習モデルによる伴奏生成

本実験では、一般的に広く知られている音楽ジャンル毎に学習曲を用意し、構築したジャンル学習モデルを用いた伴奏を生成する。ジャズ、ダンス、ロックの3つのジャンルの楽曲をそれぞれ5曲ずつ学習曲として用意し、ジャンル毎に伴奏生成モデルを構築した。このとき、各生成モデルに対する入力として同一の楽曲『大きな古時計』のメロディ、コード、および、セクション情報を用いて伴奏を生成した。図5、6、7に各ジャンルの伴奏生成モデルによって出力された伴奏の一例を示す。同一の入力情報を用いたにも関わらず、各生成モデルによって異なる伴奏が出力されていることが確認できる。

被験者25人（内12人音楽経験者）に対して、出力された伴奏のジャンル分類実験を行っ

図5 ジャズ楽曲を学習した伴奏生成モデルによる出力例

図6 ダンス楽曲を学習した伴奏生成モデルによる出力例

図7 ロック楽曲を学習した伴奏生成モデルによる出力例

表 2 ジャンルの一致率

		被験者の回答				一致率
		A	B	C	other	
学習曲 の ジャンル	A(Jazz)	16	1	5	3	0.64
	B(Rock)	0	22	2	1	0.88
	C(Dance)	0	5	19	1	0.76

表 3 印象の一致率

		被験者の回答			一致率
		激しい	明るい	どちらでもない	
学習曲 の印象	激しい	10	6	0	0.63
	明るい	3	13	0	0.81

た．実験ではまず，モデル構築に用いた学習曲に含まれないサンプル曲を 3 つのジャンル毎にそれぞれ用意し被験者に呈示した．このとき，ジャンル名は明記せず「ジャンル A」，「ジャンル B」，「ジャンル C」と表記した．被験者は提案システムが出力した楽曲とサンプル曲を聞き比べ，出力された楽曲が「ジャンル A」，「ジャンル B」，「ジャンル C」，「いずれにも相当しない (other)」のうち，どれに当てはまるかを選択する．

表 2 にジャンル分類実験の結果を示す．学習曲に用いたジャンルと被験者が回答したジャンルとの一致率について κ 係数を算出した結果， $\kappa = 0.65$ が得られた．一般的に κ 係数が 0.60 以上であるとき非常に高い一致率を意味しており，このことから，提案システムによってジャンル毎のニュアンスが十分に付与された楽曲を出力可能であることが示唆された．

5.2 印象学習モデルによる伴奏生成

本実験では，被験者が各々感じる「明るい」や「激しい」などの印象毎に学習曲を用意し，構築した印象学習モデルを用いた伴奏生成を行なった．被験者 16 名（内音楽経験者 1 名）にそれぞれ「明るい」と「激しい」という音楽の雰囲気を表す二つの印象に関して，「YAMAHA 音楽データショップ」[17] から選出した 100 曲の中から被験者自身に学習曲を 5 曲ずつ選出してもらい，選ばれた楽曲を学習曲として被験者，印象毎に伴奏生成モデルを構築した．このとき，印象学習モデルへの入力にはジャンル学習モデルによる伴奏生成実験と同一楽曲の情報を用いた．

被験者 16 名に対して，それぞれが選出した学習曲を基に構築された生成モデルの出力曲について主観評価実験を行なった．実験では，被験者はシステムが出力した伴奏付楽曲を聴取し「激しい曲」，「明るい曲」，および「どちらでもない」のどれに分類されるかを回答する．表 3 に学習曲の印象と被験者の回答の一致率を示す．学習曲の印象と被験者の回答の一致率について κ 係数を算出した結果， $\kappa = 0.52$ が得られ，比較的高い値が確認された．このことから，提案システムは被験者ごとの感性的要求や嗜好などにも応じた伴奏が生成可能であることが示唆された．

5.3 提案システムと既存システムの比較実験

提案システムと既存システム [16] の出力楽曲に対する主観評価について比較実験を行なった．既存システムは，Hidden Markov Model (HMM) を用いて曲の学習を行っており，メロディとコードをシステムへ入力することで，学習曲のニュアンスをもつベースを伴奏として入力メロディに付与した楽曲を出力する．既存システムでは生成する伴奏にベースのみを採用しているため，公正な比較を行うために本実験では提案システムの生成曲からドラムパートを取り除いた．

5.1 節で使用したジャンル毎の学習曲と入力情報を既存システムに用いて，それぞれ楽曲を生成した．本実験では，25 人の被験者（5.1 節での被験者と同一）が 5.1 節で出力された楽曲と既存システムによって出力された楽曲とを聴取し，以下の 4 つの形容詞対に関してそれぞれ 5 段階（+2，+1，0，-1，-2）で評価を行った．

好き - 嫌い

自然である - 不自然である

独創性がある - 独創性がない

快い - 不快である

システムの出力楽曲に対する上記の形容詞対についての主観評価値の平均をジャンル毎に表 4，5，6 にそれぞれ示す．Scheffe 法（有意水準 1%）による有意差検定を行った結果，形容詞対「好き - 嫌い」，「自然である - 不自然である」，「快い - 不快である」について，提案システムは既存システムに対して良好な結果を示した．このことから，提案システムは既存システムに比べ，より音楽的であり好まれる楽曲が生成可能であることが示唆された．

6. おわりに

本稿では，楽譜コンテキストのベイジアンマイニングを用いた自動伴奏付与システムを提案した．伴奏生成実験および主観評価実験によって，提案システムは学習に用いた楽曲のジャンルのニュアンス，感性的ニュアンスが付与された楽曲を出力可能であること確認し，また提案システムは既存の伴奏付与システムに対しても有意に良好な感性評価値が得られ

表 4 ジャズ楽曲学習に関する比較

	提案システム	既存システム	有意差
好き - 嫌い	0.52	-1	✓
自然である - 不自然である	0.36	-1.56	✓
独創性がある - 独創性がない	0.24	0.52	
快い - 不快である	0.52	-1.28	✓

表 5 ダンス楽曲学習に関する比較

	提案システム	既存システム	有意差
好き - 嫌い	0.64	-0.2	✓
自然である - 不自然である	-0.4	-0.52	
独創性がある - 独創性がない	1.08	-0.32	✓
快い - 不快である	0.32	-0.72	✓

表 6 ロック楽曲学習に関する比較

	提案システム	既存システム	有意差
好き - 嫌い	0.28	-0.6	✓
自然である - 不自然である	0.32	-1.28	✓
独創性がある - 独創性がない	-0.04	0.4	
快い - 不快である	0.16	-0.84	✓

る楽曲が出力可能であることを確認した。

今後の課題として、「システムの実用性向上」と「扱う楽器の種類増加」の2点が挙げられる。提案システムでは、メロディとコード、セクションの情報が必要であるが、コードとセクションの情報を自動的に認識し、メロディのみを入力とするシステムへと改良を行っていく。また、本稿ではベースとドラムのみを伴奏楽器として扱ったが、ヴァイオリンやギターなどのより多くの楽器を扱う伴奏付システムを提案する。

謝 辞

本研究は、一部、文部科学省科学研究費補助金（課題番号 20700199）、および、公益財団法人堀科学芸術振興財団の助成のもと行われた。

参 考 文 献

- 1) Hörnel, D.: "A multi-scale Neural-Network model for Learning and Reproducing Chorale Variations," *Computing in Musicology*, vol.11, pp.141-158, 1998.
- 2) Nettheim, N.: "Melodic Pattern-Detection Using MuSearch in Schubert's Die schöne Müllerin," *Computing in Musicology*, vol.11, pp.159-168, 1998.
- 3) Cope, D.: *Computers and musical style (Computer Music and Digital Audio*

series), A-R Editions, 1991.

- 4) Balaban, M., Ebcioğlu, K., Laske, O.: *Understanding Music with AI: Perspectives on Music Cognition*, AAAI Press, 1992.
- 5) 深山覚, 中妻啓, 米林裕一郎, 酒向慎司, 西本卓也, 小野順貴, 嵯峨山茂樹, "Orpheus: 歌詞の韻律に基づいた自動作曲システム (自動作曲)", 情報処理学会研究報告 [音楽情報科学], vol.2008, no.78, pp.179-184, 2008.
- 6) Manuel Alfonseca, Manuel Cebrian, Alfonso Ortega, "A simple genetic algorithm for music generation by means of algorithmic information theory", *2007 IEEE Congress on Evolutionary Computation*, pp.3035-3042, 2007.
- 7) Cubase: <http://www.steinberg.net/en/home.html>
- 8) Sonar: <http://www.cakewalk.com>
- 9) Live: <http://www.ableton.com>
- 10) Mayer, L. B.: *Emotion and meaning in Music*, University of Chicago Press, 1956.
- 11) 岡村亮吾, 山西良典, 加藤昇平, "楽譜コンテキストのページアンマイニングに基づく楽曲生成", 第73回情報処理学会全国大会, 2R-2, 東京工業大学, 2011.
- 12) Cooper, G.F., Herskovits, E.: "A bayesian method for the induction of probabilistic networks from data," *MACHINE LEARNING*, vol.9, no.4, pp.309-347, 1991.
- 13) Nattiez, J.-J., Abbate, C.: *Music and Discourse: Toward a Semiology of Music / Translated from French* Princeton Univ Press, 1990.
- 14) Stein, D.: *Engaging Music: Essays in Music Analysis*, Oxford University Press, 2004.
- 15) Ewer, G.: *Essential Chord Progressions*, Pantomime Music Publications, 2006.
- 16) Yamanishi, R., Akita, K., Kato, S.: "Automated Composing System for Submelody Using HMM: A Support System for Composing Music," *Entertainment Computing - ICEC 2010*, pp.425-427, 2010.
- 17) YAMAHA 音楽データショップ: <http://www.music-eclub.com/musicdata/>