



香川高専

卒業研究論文

論文題目

ギタータブ譜からの
ギターリフ抽出アルゴリズム

提出年月日	令和5年2月22日
学 科	電気情報工学科
氏 名	唐渡 昂希 印
指導教員（主査）	北村 大地 講師 印
副 査	柿元 健 准教授 印
学 科 長	辻 正敏 教授 印

香川高等専門学校

Guitar Riff Extraction Algorithm from Guitar Tablatures

Koki Karato

Department of Electrical and Computer Engineering
National Institute of Technology, Kagawa College

Abstract

The elements of music include melody, harmony, and rhythm. In rock music, a kind of melody, called guitar riff, has been treated as the symbol of the song. A guitar riff is a short, repeated, memorable melody, and has great influence on the impression of the song in various genres of rock music. Famous artists have developed artistic guitar riffs, and there have been trends in guitar riffs throughout each period. Therefore, we could consider guitar riffs as one of the clues to understand the style and atmosphere of music of each period. The purpose of this study is to analyze historical changes in popular music from the perspective of guitar riffs using deep neural network (DNN). The period-wise guitar riffs are used for training the DNN. This enables us to analyze the relationships of the guitar riffs in a low-dimensional latent feature space. In addition, by using such latent space, we can generate new guitar riffs with the features of a certain period or the fusion of multiple periods, which could contribute to the growth of musical art culture. To achieve this, a large number of guitar riffs is required for training the DNN. However, there is no available dataset specifically for guitar riffs, and only the dataset of the entire electric guitar score (guitar tablature) is available. To produce a dataset of guitar riffs, in this thesis, I propose new algorithms that extract guitar riffs from entire guitar tablatures. I build pattern extraction algorithms using cosine similarity or nonnegative matrix factorization. These approaches are based on an assumption that guitar riffs are repeating melody patterns. Experiments using guitar tablatures of popular songs in rock music show that the proposed method achieved an extraction accuracy of 79.7% under ideal conditions.

Keywords: guitar riff, rock music, guitar tablature

(和訳)

音楽の構成要素として、旋律、ハーモニー、及びリズム等が挙げられる。特に、ロック音楽においては、エレクトリックギターの短旋律ともいえるギターリフがその曲のシンボルとして重要視されてきた。ギターリフとは、曲中に何度も繰り返され記憶に残るエレクトリックギターの短いフレーズであり、ロック音楽を中心とした様々なジャンルにおいてその曲の印象を左右する。また、影響力のあるアーティストらによって芸術的なギターリフが次々と登場し、時代の流れとともに変遷してきた歴史がある。従って、ギターリフはいずれも、その曲が登場した年代の曲調や雰囲気等を象徴する特徴量の一つと考えられる。本研究では、そのような大衆音楽の歴史的変遷の学習と解析を目的として、深層ニューラルネットワーク (DNN) を用いて年代ごとのギターリフを学習し、低次元な潜在特徴量空間での関係性を解析することを目指す。このような潜在特徴量空間の学習によって、その年代の特徴を持つ新しいギターリフや複数の年代を融合したギターリフ等を生成できるため、音楽という誰しもが親しむ芸術文化の興隆につながると考えられる。しかしながら、DNN を学習するためには膨大な学習データが必要となるが、現状においてギターリフのデータセットは一般に提供されていない。ただし、エレクトリックギターの曲全体の楽譜 (ギタータブ譜) のデータセットは公開されているものが存在する。そこで本論文では、前述の目的を実現するための基礎研究として、曲全体のギタータブ譜からギターリフに該当する部分のみを推定し抽出するアルゴリズムについて検討する。提案手法では、ギターリフが曲中で何度も繰り返される類似旋律であることに着目し、コサイン類似度や非負値行列因子分解によるパターン抽出アルゴリズムを構築する。実際の楽曲のギタータブ譜を用いた実験では、提案手法は理想的な条件において 79.7% の抽出精度が得られることを確認した。

目次

第 1 章	緒言	1
1.1	本論文の背景	1
1.2	本論文の目的	3
1.3	本論文の構成	4
第 2 章	ギターの基礎知識	5
2.1	まえがき	5
2.2	エレキギターとタブ譜	5
2.3	ギターリフとバッキング	7
2.4	本章のまとめ	9
第 3 章	対象データと基礎技術	15
3.1	まえがき	15
3.2	DadaGP	15
3.2.1	DadaGP の概要	15
3.2.2	DadaGP 形式	15
3.3	コサイン類似度	17
3.4	NMF	17
3.4.1	NMF の概要	18
3.4.2	NMF の定式化	19
3.5	本章のまとめ	21
第 4 章	提案手法	22
4.1	まえがき	22
4.2	ギタータブ譜の数値化	22
4.2.1	ギタータブ譜の最小音価に基づく圧縮	22
4.2.2	ギタータブ譜の行列化	24
4.2.3	タブ譜行列のベクトル化	27
4.2.4	タブ譜行列の類似度行列	27
4.3	MRME 法	29
4.4	CRME 法	32

4.5	NRME 法	33
4.6	本章のまとめ	34
第 5 章	ギターリフ抽出実験	35
5.1	まえがき	35
5.2	実験条件	35
5.3	MRME 法の最適パラメータ探索と最高性能の調査	37
5.4	CRME 法の最適パラメータ探索と最高性能の調査	40
5.5	NRME 法の最適パラメータ探索と最高性能の調査	41
5.6	各提案手法の考察	44
5.7	本章のまとめ	45
第 6 章	結言	46
	謝辞	47
	参考文献	47

第 1 章

緒言

1.1 本論文の背景

エレクトリックギター（以後、エレキギターと呼ぶ）は 2010 年頃までの大衆音楽において頻繁に使用されていた。特に、ロック音楽において、歪んだ音を大音量で鳴らすことができるエレキギターの存在は必要不可欠であった。一般に、音楽の構成要素として、旋律、ハーモニー、及びリズム等が挙げられる。特に、ロック音楽においては、エレキギターの短旋律ともいえるギターリフがその曲のシンボルとして重要視されてきた。ギターリフとは、曲中に何度も繰り返され記憶に残るエレキギターの短いフレーズであり、ロック音楽を中心とした様々なジャンルにおいてその曲の印象を左右する。これまでロック音楽におけるエレキギターの演奏は、影響力のあるアーティストらによって、テクニック、スケール、及びメロディライン等の面で流行り・廃りを繰り返してきた歴史がある。エレキギターにおける重要なパートのギターリフも、そのような影響を強く受けており、その時代を象徴するギターリフのパターンが広く形成されてきたといえる。

ロック音楽史 [1, 2] によれば、ロック音楽は次に示すように発展を遂げてきた。1950 年代初頭はアメリカ南部においてリズムアンドブルースやカントリーなどの音楽ジャンルを起源としてロック音楽が形成され、1960 年代はイギリスやアメリカを中心にロック音楽が爆発的に普及した。1960 年代後半から 1970 年代にかけては、ロック音楽の黄金時代とも呼ばれ、ラテンロック・カントリーロック・ブルースロック・ハードロック・プログレッシブロック等の様々なジャンルへと派生し、バラエティ豊かなギターリフが登場した。その後、1970 年代後半では産業的になりすぎたロック音楽を批判するように、再びストレートでシンプルな曲調（いわゆるパンクロック）に回帰する動きがあり、やはり当時のギターリフにもその社会的な雰囲気強く反映されている。その後の 1980 年代は、当時の人々が抱える何者かへの怒りを直接的に表現したオルタナティブロックやグランジロックが登場し、重く激しいロック音楽が台頭した。1980 年代後半から 1990 年代にかけては再びファンクやヒップホップ等と融合したバラエティ豊かなロック音楽が登場した他、黄金時代の影響を受けた王道のロック音楽を再現するアーティストも人気を博した。1990 年代以降は、ポストロックと呼ばれる多様な音楽スタイルへと分岐・発展していくため、ジャンルとして分類すること自体が難しくなっていく。

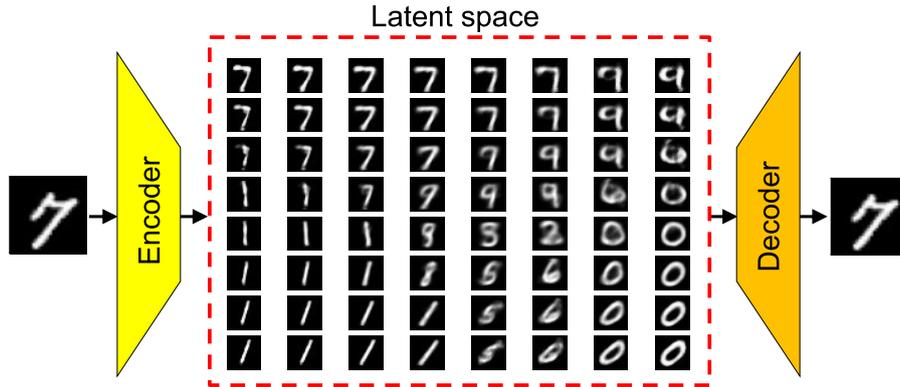


Fig. 1.1. Overview of VAE model.

しかしながら、1990年代から2000年代にかけてのギターリフは音色や響きを重視したものが多く登場しており、やはり何らかの特徴は時代と結びついている可能性がある。以上のことから、様々な年代のギターリフの特徴に着目して解析することは、ロック音楽を含む大衆音楽の歴史の変遷を理解する一つ方法になると考えられる。

一方、近年では深層学習の登場もあり、自動作曲を主目的とした技術及び手法が新しく登場している。その中でも、本研究の最終的な目的とよく類似した手法として、深層学習モデルの一つである変分自己符号化器 (variational autoencoder: VAE) [3, 4] を用いた MusicVAE [5] について述べる。まず、VAE とは、次元圧縮した低次元特徴量がある確率分布に従うという仮定の下、入力と同じデータを出力するような (低次元の) 潜在特徴量空間を学習する深層ニューラルネットワークである。これによって、潜在空間及び潜在特徴量そのものに対して確率分布に基づく一定の解釈性を持たせることができ、データを低次元な特徴量空間で分析することが可能となる。Fig. 1.1 は手書き数字画像を VAE で学習した例を表している。Fig. 1.1 中の中央に示す画像は、潜在空間における低次元の特徴量をデコーダに通したときに生成される画像を、特徴量の座標毎に示したものである。潜在空間には、複数クラス (この例では手書き数字の種類) の間を補完するような画像 (例えば「7」と「9」の間の画像等) の特徴量が存在していることが分かる。このような潜在特徴量空間の学習を膨大な楽譜データに対して適用したものが MusicVAE である。つまり、MusicVAE は複数の楽曲にまたがる特徴量を低次元の潜在空間に埋め込んでおり、潜在空間上で任意の乱数を生成しデコーダを通せば、その特徴量を持った新たな楽譜を生成できる。この MusicVAE の全体像を Fig. 1.2 に示す。エンコーダに、双方向再帰型ニューラルネットワーク (recurrent neural network: RNN) を用いることで、エンコーダが楽譜の繰り返し構造を学習している。デコーダには、階層構造の RNN を用いることで、潜在空間が楽譜の大域的な構造と局所的な構造をそれぞれ学習できるように工夫されている。これによって、MusicVAE は複数の学習データの特徴を合わせ持つ、新たな旋律を生成することに成功している。MusicVAE のように、ギターリフの特徴量を、年代という概念を新たに付与しながら学習させることができれば、ロック音楽の歴史の変遷を知る一つの手掛かりとなると考える。

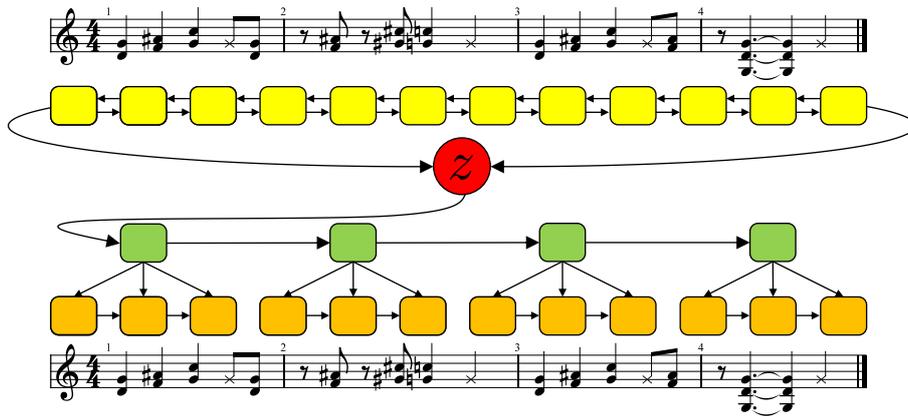


Fig. 1.2. Overview of MusicVAE model using hierarchical recurrent VAE [5].

1.2 本論文の目的

本研究の最終的な目的は、ギターリフの特徴量解析を行うことで、ロック音楽の歴史の変遷を可視化すること及び各時代の特徴量を持った、あるいは融合したギターリフを生成することである。ここで特徴量とは、何らかの次元圧縮手法で求めることのできる低次元空間の潜在的なものであり、理想的には年代のラベルが付いた学習データから深層学習で得られる「その時代らしさ」や「その時代の雰囲気」等を含むものである。この年代のように、各データに何らかのラベルが付与された学習データから、条件付きの特徴量を学習するには、条件付き変分自己符号化器 (conditional variational autoencoder: CVAE) [6] を利用する方法が挙げられる。CVAE は、VAE のエンコーダ及びデコーダにラベルの情報を加えて潜在空間を学習することで、同じラベルを持つデータの潜在特徴量が近接的に分布する様に学習することができる。本研究では、Fig. 1.3 に示す様に、ギターリフ及びその楽曲のリリース年代を学習させる。これによって、ギターリフの年代に対する変遷等が潜在特徴量空間に埋め込まれ、結果的に歴史を解析することができるようになることを期待できる。また、このような潜在空間を得ることで、次に示すような応用が期待できる。

- 各ギターリフの潜在空間上の距離から、アーティスト間の（影響を受ける・与えるといった）繋がりを確認できる
- 潜在空間上に乱数を与えて CVAE のデコーダを通すことで、その年代のギターリフの特徴量を反映した新しいギターリフの生成や、複数の年代の特徴量を融合したギターリフの生成ができる
- 典型的なギターリフの構造が潜在特徴量空間に埋め込まれるため、多くのギターリフに見られる共通のテクニックやパターン等を知ることができ、エレキギターの演奏法の教育や教材に利用できる

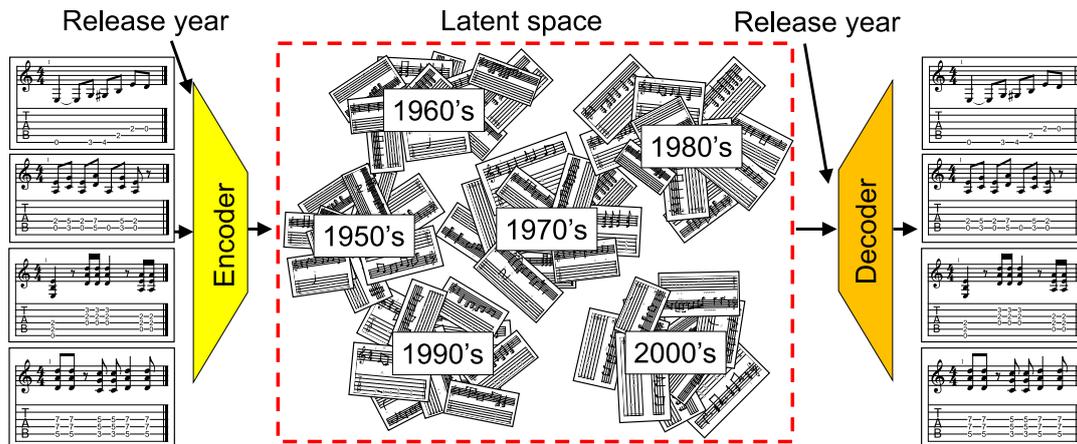


Fig. 1.3. Overview of final goal of this research topic.

本研究を行う上で、CVAEの潜在空間を学習するためには、年代のラベルが付与された大量のギターリフのギタータブ譜が必要となる。しかし、現状において、曲全体を含むギタータブ譜のデータセットは存在するが、ギターリフの部分のみを含むギタータブ譜のデータセットは存在しない。そこで、本論文ではギタータブ譜からギターリフのみを抽出する手法を新たに提案する。本論文では、ギターリフが曲の中で繰り返し演奏される類似パターンであることに着目して、ギターリフを抽出するアルゴリズムを構築する。実験では、ロック音楽史に残る有名な楽曲のギタータブ譜に対して提案手法を適用し、どの程度の精度で正解のギターリフを抽出できるかについて調査する。

1.3 本論文の構成

2章では、ロック音楽において、必要不可欠な存在となるギターについて述べる。また、ギター、ベースなどの弦楽器において使用されるタブ譜について述べ、本研究の要であるギターリフについて述べる。3章では、本論文で使用するギターやベース等の弦楽器のタブ譜のデータセットである DadaGP [7] について述べる。本論文では、DadaGP によって提供されるデータを何らかの形で数値に変換し、ギターリフを抽出する。そこで用いられるギターリフの特徴として、ギターリフは何度も繰り返される旋律であることが挙げられる。その繰り返しパターンを認識するための数理理論の基礎知識として、コサイン類似度及び非負値行列因子分解 (nonnegative matrix factorization: NMF) [8, 9] について述べる。4章では、DadaGP によって提供されるデータを行列及びベクトルに変換する方法について述べる。また、提案手法である MRME 法 (measure-based repeated melody extraction), CRME 法 (cosine-similarity-based repeated melody extraction), 及び NRME 法 (NMF-based repeated melody extraction) についても述べる。5章では、提案手法を用いてギターリフ抽出実験を行い、どの程度の精度で正しくギターリフを抽出できるかを確かめる。6章では、本論文の結果についてまとめ、今後の課題を述べる。

第 2 章

ギターの基礎知識

2.1 まえがき

本章では、ロック音楽において必要不可欠となるギターについて述べる。2.2 節では、エレキギターについて述べ、弦楽器において頻繁に使用されるタブ譜について述べる。2.3 節では、ギターリフについて述べ、歴史的に有名なギターリフの例をいくつか挙げる。さらに、曲の中で何度も繰り返されるという、ギターリフと同じ特徴を持つ、バックングについて述べる。2.4 節で本章についてまとめる。

2.2 エレキギターとタブ譜

エレキギターの外観を Fig. 2.1 に示す。通常のエレキギターは、押さえる弦として第 1 弦から第 6 弦あり、各弦の押さえる位置として第 0 フレットから第 22 フレットまでであることが多い。ただし、第 0 フレットは弦を押さえることなく弾くことになる。

タブ譜とは、ギター、ベース、及びウクレレ等のフレットを持つ弦楽器において、一般的な五線譜の音価表記と弦毎の運指（押下する位置）を組み合わせた楽譜であり、特にギターの楽譜としては広く普及している表記法である。ギター用のタブ譜（以下、ギタータブ譜と呼ぶ）とベース用のタブ譜の例を Figs. 2.2 及び 2.3 にそれぞれ示す。タブ譜は、その楽器の弦の本数だけ水平線が引かれており、ギタータブ譜であれば通常は 6 本線となる。この水平線上に、4 分音符や 8 分休符等の五線譜と同じ音価表記を持つ音符が表記されるが、各音符には押下する弦と位置（フレット）も記載されている。従って奏者は、指定された音符のタイミングで指定された弦・位置を押下しながら弦を弾くことが求められる。このような専用の楽譜が使用される理由として、ギターやベース等の一般的な弦楽器は同じ音高を発音する方法が複数存在することが挙げられる。例えば、通常の調弦（いわゆるスタンダードチューニング、第 6 弦から第 1 弦まで順に、開放弦の音高が E2, A2, D3, G3, B3, E4 となる）において、第 6 弦・第 12 フレット、第 5 弦・第 7 フレット、及び第 4 弦・第 2 フレットは全て同じ E3 という音高である。従って、通常の五線譜ではどのように弦を押下するかが決まらない問題が生じる。すな

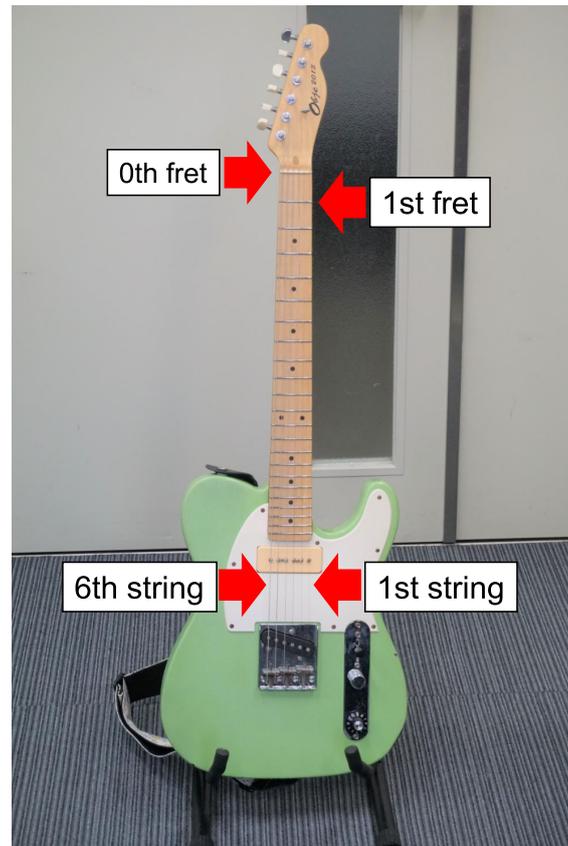


Fig. 2.1. Strings and frets of electric guitar.

わち、タブ譜は五線譜に対して演奏の運指の情報を追加した楽譜と捉えられる。さらに、その楽器ならではのバラエティに富んだ奏法も同時に記載されている。ギタータブ譜では、一例として下記の奏法記述が存在する。

- ビブラート
- チョーキング (ベンディング)
- ブラッシング
- ハンマリングオン
- スライド
- ブリッジミュート (パームミュート)
- ハーモニクス

上記の例のギタータブ譜での表記を Fig. 2.4 に示す。奏者はこれらの演奏法に忠実に従うことで、可能な限り作曲者の意図した音を再現することができる。

(I Can't Get No) Satisfaction
The Rolling Stones
 Out of Our Heads

Standard tuning
 ♩ = 138

od.guit.

Fig. 2.2. Guitar tablature of song “(I Can’t Get No) Satisfaction” by The Rolling Stones [7].

(I Can't Get No) Satisfaction
The Rolling Stones
 Out of Our Heads

Standard tuning
 ♩ = 138

el.bs.

Fig. 2.3. Bass tablature of song “(I Can’t Get No) Satisfaction” by The Rolling Stones [7].

Vibrato Bending Brushing Hammering-on Pulling-off Palm mute Harmonics

Fig. 2.4. Examples of notations in guitar tablature for representing techniques and articulations.

2.3 ギターリフとバックイング

ギターリフとは、曲の要所で繰り返し登場するエレキギターの短い旋律のことであり、1小節から8小節程度の長さであることが多い。イントロで繰り返し使われることが多いが、間奏やサビ（コーラス）等で使われることもある。典型的なギターリフの曲中での繰り返しパターンを Fig. 2.5 に示す。ここで、長方形の枠は1小節を表し、アルファベットはギターリフの旋律、ギリシャ文字はギターリフではない旋律を表す。

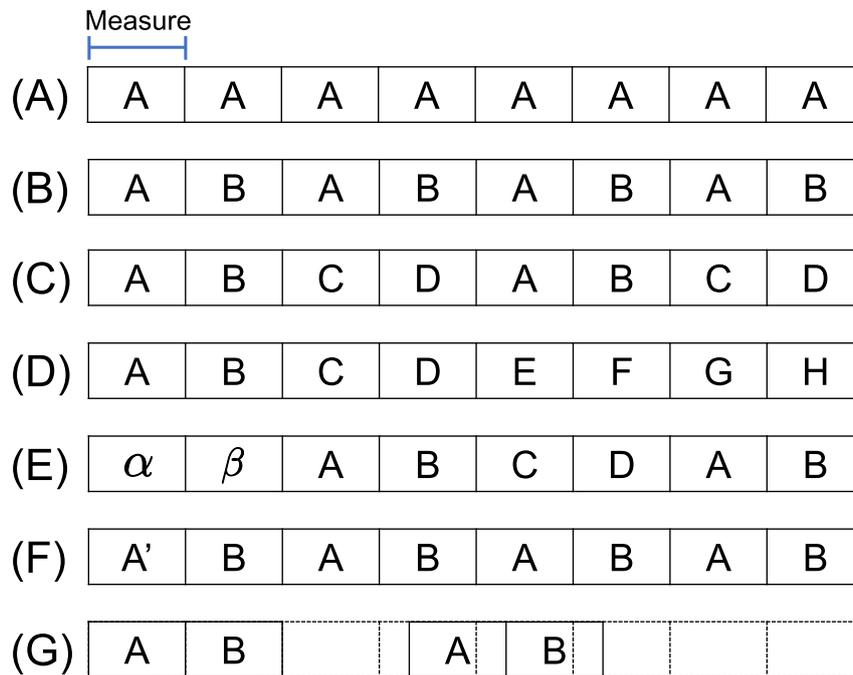


Fig. 2.5. Typical repetition patterns of guitar riff.

- (A) に該当するギターリフの一例を Fig. 2.6 に示す。これは、Wild Cherry の “Play That Funky Music” を一部抜粋したもので、1 小節目がギターリフである。
- (B) に該当するギターリフの一例を Fig. 2.7 に示す。これは、Marilyn Manson の “Sweet Dreams” を一部抜粋したもので、1~2 小節目がギターリフである。
- (C) に該当するギターリフの一例を Fig. 2.8 に示す。これは、Deep Purple の “Smoke on the Water” を一部抜粋したもので、1~4 小節目がギターリフである。ギタータブ譜上では、このリフが冒頭から 6 回繰り返される。
- (D) に該当するギターリフの一例を Fig. 2.9 に示す。これは、Iron Maiden の “Fear of the Dark” を一部抜粋したもので、1~8 小節目がギターリフである。
- (E) に該当するギターリフの一例を Fig. 2.10 に示す。これは、AC/DC の “Back in Black” を一部抜粋したものである。ギターリフは必ず 1 小節目から始まるわけではなく、この楽曲におけるギターリフは 3~6 小節目となる。
- (F) に該当するギターリフの一例を Fig. 2.11 に示す。これは、ZZ Top の “Sharp Dressed Man” を一部抜粋したもので、1~2 小節目がギターリフである。1 小節目及び 3 小節目の冒頭のリズムが少し異なっている。この例から分かるように、ギターリフは完全に同じ旋律が繰り返されるわけではない。
- (G) に該当するギターリフの一例を Fig. 2.12 に示す。これは、Led Zeppelin の “Black Dog” を一部抜粋したもので、6~11 小節目がギターリフである。ギターリフは繰り返されるものだが、小節による旋律の区切り位置が必ず同じになるとは限らないことを示す稀な例である。

エレキギターの演奏において頻繁に繰り返されるもう一つのパターンとして、バックキングが挙げられる。バックキングとは、同じ音の組み合わせだけを何度も演奏するギターの演奏パートのことであり、主に和音、パワーコード、オクターブコード等を単調に繰り返すものである。

バックキングの一例を Fig. 2.13 に示す。Fig. 2.13 のギタータブ譜は Blur の “Song 2” という曲におけるバックキングの冒頭である。Fig. 2.13 では、冒頭の 2 小節のみを示しているが、同じ旋律を曲全体を通して弾き続ける。その他にも、Fig. 2.14 のような弦 3 本を用いるバックキング、Fig. 2.15 のような弦 5 本を用いるバックキング、及び Fig. 2.16 のような弦 1 本を用いるバックキングも存在する。バックキング奏法はほとんどの楽曲に含まれており、その曲の特徴を強く表すものではないため、(線引きは多少曖昧ではあるが) ギターリフには含めないことが一般的である。

2.4 本章のまとめ

本章では、ギターに関する基礎知識として、エレキギター及び弦楽器において頻繁に使用される楽譜であるタブ譜について述べた。また、本研究の抽出対象であるギターリフについて述べ、似た特徴を持つバックキングについて述べた。次章では、本論文で使用するギターやベース等の弦楽器のタブ譜のデータセットである DadaGP [7] について述べる。また、ギターリフが繰り返されていることを認識するための数理理論の基礎知識として、コサイン類似度及び NMF [8, 9] について述べる。

Play That Funky Music
Wild Cherry
 Wild Cherry

Standard tuning
 ♩ = 110

Fig. 2.6. Lead guitar tablature of song “Play That Funky Music” by Wild Cherry [7].

Sweet Dreams
Marilyn Manson
 Smells Like Children

Standard tuning
 ♩ = 90

Fig. 2.7. Lead guitar tablature of song “Sweet Dreams” by Marilyn Manson [7].

Smoke on the Water
Deep Purple
 Machine Head

Standard tuning
 ♩ = 112

Fig. 2.8. Lead guitar tablature of song “Smoke on the Water” by Deep Purple [7].

Fear of the Dark
Iron Maiden
 Fear of the Dark

Standard tuning
 ♩ = 80

Fig. 2.9. Lead guitar tablature of song “Fear of the Dark” by Iron Maiden [7].

Back in Black
AC/DC
Back in Black

Standard tuning
♩ = 96

Fig. 2.10. Lead guitar tablature of song “Back in Black” by AC/DC [7].

Sharp Dressed Man
ZZ Top
Eliminator

Standard tuning
♩ = 124

Fig. 2.11. Lead guitar tablature of song “Sharp Dressed Man” by ZZ Top [7].

Black Dog
Led Zeppelin
Led Zeppelin IV

Standard tuning
♩ = 150

od.guit.

Fig. 2.12. Lead guitar tablature of song “Black Dog” by Led Zeppelin [7].

Song 2
Blur
Blur

Standard tuning
♩ = 130

dis.guit.

Fig. 2.13. Backing guitar tablature of song “Song 2” by Blur [7].

Mr. Brownstone
Guns N' Roses
Appetite for Destruction

Tune down 1/2 step
① = E♭ ④ = D♭
② = B♭ ⑤ = A♭
③ = G♭ ⑥ = E♭

♩ = 105

od.guit.

Fig. 2.14. Backing guitar tablature of song “Mr. Brownstone” by Guns N’ Roses [7].

From Me to You
The Beatles
 Standard tuning Twist and Shout
 ♩ = 140

jz.guit.

Fig. 2.15. Backing guitar tablature of song “From Me to You” by The Beatles [7].

Eye of the Tiger
Survivor
 Standard tuning
 ♩ = 112

s.guit.

Fig. 2.16. Backing guitar tablature of song “Eye of the Tiger” by Survivor [7].

第 3 章

対象データと基礎技術

3.1 まえがき

本章では、本論文で使用するデータセット及び提案手法で使用する基礎知識について述べる。3.2 節ではギターやベース等の弦楽器のタブ譜のデータセットである DadaGP [7] について述べる。3.3 節ではコサイン類似度について述べる。この技術は、ギターリフ抽出を行う際に、ギターリフが繰り返される旋律であることを認識するためのものである。3.4 節では NMF [8, 9] について述べる。NMF はある非負行列に含まれるパターンを低ランク近似で抽出する手法と解釈できる。これにより、ギタータブ譜を非負行列で表すことができるならば、パターン抽出を用いてギタータブ譜からギターリフを抽出するところができると考える。3.5 節で本章についてまとめる。

3.2 DadaGP

3.2.1 DadaGP の概要

DadaGP [7] とは、ギターやベース等の弦楽器のタブ譜 26,181 曲を提供するデータセットのことである。DadaGP では、各楽器のタブ譜が、Guitar Pro [10] と呼ばれるタブ譜用編集ソフトウェアで開くことのできるファイル（以後、Guitar Pro 形式と呼ぶ）及び独自に定義された形式（以後、DadaGP 形式と呼ぶ）でタブ譜情報が記述されたテキストファイルの 2 種類で提供されている。2.3 節で示したギタータブ譜は全て DadaGP によって提供されたものである。また、DadaGP では Guitar Pro 形式及び DadaGP 形式を相互に変換するエンコーダ及びデコーダの Python で記述されたスクリプトも提供されている。

3.2.2 DadaGP 形式

本項では、DadaGP 形式におけるタブ譜の記述方法について説明する。GuitarPro 形式で記述されたタブ譜及びそれに対応する DadaGP 形式のテキストデータを Fig. 3.1 に示す。全

Standard tuning

♩ = 190

(a)

(b)

```

01: artist:example
02: downtune:0
03: tempo:190
04: start
05: new_measure
06: distorted0:note:s5:f0
07: nfx:palm_mute
08: wait:480
09: distorted0:note:s5:f7
10: nfx:hammer
11: wait:240
12: distorted0:note:s5:f5
13: nfx:hammer
14: wait:240
15: distorted0:note:s5:f10
16: nfx:bend:type1
17: wait:960
18: distorted0:note:s5:f5
19: wait:1920
20: end

```

Fig. 3.1. Example of (a) guitar tablature (Guitar Pro format) and (b) its DadaGP format file.

この曲は `artist`, `downtune`, 及び `tempo` が指定されている。これは曲のメタ情報であり, `artist` がその曲のアーティスト名, `downtune` がそのギタータブ譜の調弦の種類 (0 はスタンダードチューニング, -1 は半音下げチューニング等), `tempo` はその曲の beat per minutes を表している。これらのメタ情報に続いて, 曲の開始を表す `start` が記述される。また, 各小節の開始位置には `new_measure` 及び曲の終わりには `end` が記述される。

次に運指の記述方法について説明する。一つの運指は `instrument:note:string:fret` (以降, この記述を楽器トークンと記述する) という並びで記述される。ここで `instrument` とは `distorted guitar`, `clean guitar`, 及び `bass` 等の楽器パートの種類を表し, `note` はその行が音符に関して表記していることを表す。また, `string` 及び `fret` は `s1`, `s2`, ... 及び `f0`, `f1`, ... のように記述する。例えば, `distorted0:note:s1:f0` と記述されているとする。これは `distorted0` ギターの第1弦・第0フレットに対応している。`distorted0` における0のように, 楽器パート名の後に書かれている数字は, 同一楽曲内の同一種類の楽器を区別するために振られている番号である。また, 休符の場合は, `instrument:rest` と記述する。これらの音符及び休符の長さについては `wait` によって数値的に指定する。DadaGP 形式では, 4分音符の長さを 960 tick と定めている。従って, 例えば, `distorted0` という楽器パートの第1弦・第0フレットを4分音符として鳴らす場合の表記は, `distorted0:note:s1:f0` の行の直後に `wait:960` となる。

2.2 節で述べたように, ギターやベース等には同じ音符やその並びに対して, 様々な異なる奏法が存在する。この奏法を DadaGP 形式で記述する場合は, Fig. 3.1 の DadaGP 形式

ファイルの7行目や10行目等のように楽器トークンの直後に奏法の種類を記述する。例えば、`distorted0`という楽器パートの第1弦・第1フレットの音をチョークアップ(バンドアップ)という奏法で演奏する場合には、`distorted0:note:s1:f1`の行の直後に`nfx:bend:type1`、さらにその行の直後に`wait:960`となる。ここで、`type1`は、チョークアップ(`type1`)、チョークダウン(`type2`)、チョークホールド(`type3`)、及びプリチョーク(`type4`)等を指定するためである。

3.3 コサイン類似度

本論文では、ギタータブ譜からギターリフを抽出する上で、ギタータブ譜を数値に変換し、ある2つの旋律が類似しているかどうかを求める。本節では、2つの旋律の類似度を求める際に必要となるコサイン類似度について述べる。

実ベクトル $\mathbf{x}, \mathbf{y} \in \mathbb{R}^{\tilde{N}}$ の内積 $\langle \mathbf{x}, \mathbf{y} \rangle$ を次のように定める。

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{\tilde{n}=1}^{\tilde{N}} x_{\tilde{n}} y_{\tilde{n}} \quad (3.1)$$

ここで、 $x_{\tilde{n}}$ 及び $y_{\tilde{n}}$ は、それぞれ \mathbf{x} 及び \mathbf{y} の第 \tilde{n} 成分であることを表す。次に、実ベクトル \mathbf{x}, \mathbf{y} に対するコサイン類似度 $\mathcal{S}_{\cos}(\mathbf{x}, \mathbf{y})$ を次のように定義する。

$$\mathcal{S}_{\cos}(\mathbf{x}, \mathbf{y}) = \begin{cases} \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\sqrt{\langle \mathbf{x}, \mathbf{x} \rangle \langle \mathbf{y}, \mathbf{y} \rangle}} & \langle \mathbf{x}, \mathbf{x} \rangle \langle \mathbf{y}, \mathbf{y} \rangle \neq 0 \\ 0 & \text{otherwise} \end{cases} \quad (3.2)$$

コサイン類似度は、2つのベクトルがどの程度類似しているかを表す1つの指標である。証明は省略するが、コサイン類似度に関して次式が成立する。

$$\mathcal{S}_{\cos}(\mathbf{x}, \mathbf{y}) = 1 \Leftrightarrow \mathbf{x} = \mathbf{y} \quad (3.3)$$

$$-1 \leq \mathcal{S}_{\cos}(\mathbf{x}, \mathbf{y}) \leq 1 \quad (3.4)$$

コサイン類似度の特徴的な性質として、 $\tilde{n} = 2, 3, \dots, \tilde{N}$ について $0 < x_{\tilde{n}}, y_{\tilde{n}} \ll y_1$ が成り立つとき、次式が成立する。

$$\mathcal{S}_{\cos}(\mathbf{x}, \mathbf{y}) \approx \begin{cases} 1 & x_1 \neq 0 \\ 0 & \text{otherwise} \end{cases} \quad (3.5)$$

これは、コサイン類似度が上記の場合において、特定の情報だけに依存することを示している。この特徴については、4.2.2項で、ギタータブ譜を行列に変換する際に、再び言及する。

3.4 NMF

本節では、NMF [8, 9] について述べる。3.4.1項でNMFの概要について述べた後、3.4.2項でNMFを定式化し、解を求めるための更新式について述べる。そして、NMFの誤差関数によって生じる違いについて述べる。

3.4.1 NMF の概要

NMF は与えられた非負値行列 $\mathbf{X} \in \mathbb{R}_{\geq 0}^{I \times J}$ を次式で近似するアルゴリズムである。

$$\mathbf{X} \approx \mathbf{W}\mathbf{H} \quad (3.6)$$

ここで、 $\mathbf{W} \in \mathbb{R}_{\geq 0}^{I \times K}$ 及び $\mathbf{H} \in \mathbb{R}_{\geq 0}^{K \times J}$ は NMF によって求められる行列である。NMF の特徴として、 K を十分小さくとり、低ランク行列の積によって \mathbf{X} を表現することが挙げられる。ここで、低ランク近似の例を示すために、次の非負値行列が与えられたとする。

$$\mathbf{X} = \begin{bmatrix} 1 & 3 & 4 \\ 2 & 6 & 9 \\ 3 & 9 & 10 \end{bmatrix} \quad (3.7)$$

このとき、 $K = 1$ の条件の下で NMF による行列分解を考えたとき、次のような結果があり得る。

$$\mathbf{X} \approx \mathbf{W}\mathbf{H} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} [1 \ 3 \ 4] \quad (3.8)$$

$$= \begin{bmatrix} 1 & 3 & 4 \\ 2 & 6 & 8 \\ 3 & 9 & 12 \end{bmatrix} \quad (3.9)$$

この誤差を表す行列 $\mathbf{D} = \mathbf{X} - \mathbf{W}\mathbf{H}$ は次のようになる。

$$\mathbf{D} = \mathbf{X} - \mathbf{W}\mathbf{H} \quad (3.10)$$

$$= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & -2 \end{bmatrix} \quad (3.11)$$

\mathbf{X} を低ランク行列によって精度良く近似するためには、実際に \mathbf{X} が低ランク行列であるか、フルランクであっても低ランク行列に微小な乱数がのったような状態の行列（すなわち、式 (3.11) のような誤差行列のノルムが小さな値になるような行列）である必要がある。NMF によって得られる \mathbf{W} は \mathbf{X} に頻出する列ベクトルが現れ、 \mathbf{H} はその列ベクトルの大きさを表すことになる。上記の場合において、 \mathbf{X} の 1 列目 \mathbf{x}_1 及び 2 列目 \mathbf{x}_2 が線形従属の関係 $3\mathbf{x}_1 = \mathbf{x}_2$ にある。これによって、頻出する \mathbf{x}_1 が \mathbf{W} として得られ、 \mathbf{X} の 1 列目及び 2 列目が完全に復元される。

より一般に、 $\text{rank}(\mathbf{X}) \leq \min(I, J)$ 及び $\text{rank}(\mathbf{W}\mathbf{H}) \leq \min(I, J, K)$ が成り立つ。よって、 $K \geq \min(I, J)$ を満たすように K を定めれば、NMF によって $\mathbf{X} = \mathbf{W}\mathbf{H}$ となる \mathbf{W} 及び \mathbf{H} を得ることができる（例えば \mathbf{W} が単位行列かつ $\mathbf{X} = \mathbf{H}$ 等）が、このような自明な分解からは有益な情報を得ることはできない。NMF における有意な情報の抽出には、 \mathbf{X} を低ランク行列 $\mathbf{W}\mathbf{H}$ で近似することが本質的となる。もし \mathbf{X} が低ランクであるならば $K \ll \min(I, J)$ となるように K を定めたとしても、適切な \mathbf{W} 及び \mathbf{H} は \mathbf{X} を高精度に近似することができ

るはずである。そのようなときに得られる \mathbf{W} や \mathbf{H} には、 \mathbf{X} を構成する少数の（高々 K 個の）有意な情報が基底ベクトル及びその係数として得られることになる。

具体的には、 $\mathbf{X} = [\mathbf{x}_1 \ \mathbf{x}_2 \ \cdots \ \mathbf{x}_J]$, $\mathbf{W} = [\mathbf{w}_1 \ \mathbf{w}_2 \ \cdots \ \mathbf{w}_K]$, 及び \mathbf{H} の (k, j) 成分を h_{kj} と表す時、これらの関係式は次式で表すことができる。

$$\mathbf{x}_j = \sum_{k=1}^K h_{kj} \mathbf{w}_k \quad (3.12)$$

これにより、 $K \ll \min(I, J)$ として \mathbf{X} を高精度に近似できた場合、 \mathbf{X} は高々 K 個の基底ベクトルによって構成されていることになる。従って、NMF は非負行列 \mathbf{X} に含まれる潜在的な頻出パターンを \mathbf{W} や \mathbf{H} として抽出できる教師無し学習（パターン認識）であるといえる。

3.4.2 NMF の定式化

ある非負行列 $\mathbf{X} \in \mathbb{R}_{\geq 0}^{I \times J}$ が与えられた時、NMF を次式で定式化する。

$$\min_{\mathbf{W}, \mathbf{H}} \mathcal{D}(\mathbf{X} | \mathbf{W}\mathbf{H}) \text{ s.t. } w_{ik}, h_{kj} \geq 0 \ \forall i = 1, 2, \dots, I, \ j = 1, 2, \dots, J, \ k = 1, 2, \dots, K \quad (3.13)$$

ただし、 w_{ik} 及び h_{kj} は $\mathbf{W} \in \mathbb{R}_{\geq 0}^{I \times K}$ の (i, k) 成分及び $\mathbf{H} \in \mathbb{R}_{\geq 0}^{K \times J}$ の (k, j) 成分であり、 $\mathcal{D}(\cdot)$ は同じサイズの行列の類似度を求める関数である。式 (3.13) の解は閉形式では与えられない。しかし、 \mathbf{W} 及び \mathbf{H} を適当な非負乱数で初期化し、反復計算によって解を求めることができる。

本論文では、行列の類似度を求める関数として、次の3つを扱う。

- 二乗 Euclid 距離

$$\mathcal{D}_{\text{Eu}}(\mathbf{A} | \mathbf{B}) = \sum_{i=1}^I \sum_{j=1}^J (a_{ij} - b_{ij})^2 \quad (3.14)$$

- 一般化 Kullback–Leibler ダイバージェンス

$$\mathcal{D}_{\text{KL}}(\mathbf{A} | \mathbf{B}) = \sum_{i=1}^I \sum_{j=1}^J \left[a_{ij} \log \frac{a_{ij}}{b_{ij}} - (a_{ij} - b_{ij}) \right] \quad (3.15)$$

- 板倉斎藤擬距離

$$\mathcal{D}_{\text{IS}}(\mathbf{A} | \mathbf{B}) = \sum_{i=1}^I \sum_{j=1}^J \left(\frac{a_{ij}}{b_{ij}} - \log \frac{a_{ij}}{b_{ij}} - 1 \right) \quad (3.16)$$

ここで a_{ij} 及び b_{ij} は、それぞれ \mathbf{A} 及び \mathbf{B} の (i, j) 成分であることを表す。

式 (3.13) に式 (3.14) を適用すれば、更新式は次のようになる。

$$\mathbf{W} \leftarrow \mathbf{W} \odot [\mathbf{X}\mathbf{H}^T] \oslash [\mathbf{W}\mathbf{H}\mathbf{H}^T] \quad (3.17)$$

$$\mathbf{H} \leftarrow \mathbf{H} \odot [\mathbf{W}^T\mathbf{X}] \oslash [\mathbf{W}^T\mathbf{W}\mathbf{H}] \quad (3.18)$$

ここで、 \odot 及び \oslash は、それぞれ行列の要素毎の積 (Hadamard 積) 及び商を表し、 \cdot^T は転置を表す記号である。また、 \leftarrow は変数更新を表す演算子である。

式 (3.13) に式 (3.15) を適用すれば、更新式は次のようになる。

$$\mathbf{W} \leftarrow \mathbf{W} \odot [(\mathbf{X} \oslash (\mathbf{W}\mathbf{H})) \mathbf{H}^T] \oslash [\mathbf{1}^{(I \times J)} \mathbf{H}^T] \quad (3.19)$$

$$\mathbf{H} \leftarrow \mathbf{H} \odot [\mathbf{W}^T (\mathbf{X} \oslash (\mathbf{W}\mathbf{H}))] \oslash [\mathbf{W}^T \mathbf{1}^{(I \times J)}] \quad (3.20)$$

ここで、 $\mathbf{1}^{(I \times J)}$ は全ての成分が 1 である $I \times J$ の行列である。

式 (3.13) に式 (3.16) を適用すれば、更新式は次のようになる。

$$\mathbf{W} \leftarrow \mathbf{W} \odot \left[\left\{ \mathbf{X} \oslash (\mathbf{W}\mathbf{H}) \right\}^2 \mathbf{H}^T \right]^{\cdot \frac{1}{2}} \oslash \left[\left\{ \mathbf{1}^{(I \times J)} \oslash (\mathbf{W}\mathbf{H}) \right\} \mathbf{H}^T \right]^{\cdot \frac{1}{2}} \quad (3.21)$$

$$\mathbf{H} \leftarrow \mathbf{H} \odot \left[\mathbf{W}^T \left\{ \mathbf{X} \oslash (\mathbf{W}\mathbf{H}) \right\}^2 \right]^{\cdot \frac{1}{2}} \oslash \left[\mathbf{W}^T \left\{ \mathbf{1}^{(I \times J)} \oslash (\mathbf{W}\mathbf{H}) \right\} \right]^{\cdot \frac{1}{2}} \quad (3.22)$$

ここで、 \cdot^2 及び $\cdot^{\frac{1}{2}}$ は要素毎の 2 乗及び平方根を表す。

式 (3.13) は、 \mathbf{X} の各成分が確率分布に従う仮定を導入すれば、 \mathbf{W} 及び \mathbf{H} に関する最尤推定問題と等価であることが示される。以下ではこれについて述べる。

\mathbf{X} の要素 x_{ij} が互いに独立で正規分布 $N(\theta_{ij}, \sigma^2)$ に従うことを仮定する。このとき、確率密度関数を $f_N(x_{ij}|\theta_{ij}, \sigma^2)$ とすれば、対数尤度 L_N は次のように変形することができる。

$$L_N = \log \left\{ \prod_{i=1}^I \prod_{j=1}^J f_N(x_{ij}|\theta_{ij}, \sigma^2) \right\} \quad (3.23)$$

$$= \sum_{i=1}^I \sum_{j=1}^J \log (f_N(x_{ij}|\theta_{ij}, \sigma^2)) \quad (3.24)$$

$$= \sum_{i=1}^I \sum_{j=1}^J \log \left[\frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{(x_{ij} - \theta_{ij})^2}{2\sigma^2} \right\} \right] \quad (3.25)$$

$$= \sum_{i=1}^I \sum_{j=1}^J \left[-\frac{1}{2\sigma^2} (x_{ij} - \theta_{ij})^2 + \log \left(\frac{1}{\sqrt{2\pi\sigma^2}} \right) \right] \quad (3.26)$$

$$= -\alpha_N \mathcal{D}_{\text{Eu}}(\mathbf{X}|\Theta) + \beta_N \quad (3.27)$$

ここで、 $\Theta \in \mathbb{R}^{I \times J}$ は θ_{ij} を要素に持つ行列である。また、 $\alpha_N > 0$ 及び β_N は θ_{ij} に依存しない定数である。よって、 L_N を最大化することは二乗 Euclid 距離を最小化することに等しい。

次に、 \mathbf{X} の要素 x_{ij} が互いに独立でポアソン分布 $\text{Po}(\theta_{ij})$ に従うことを仮定する。このと

き，確率密度関数を $f_{\text{Po}}(x_{ij}|\theta_{ij})$ とすれば，対数尤度 L_{Po} は次のように変形することができる．

$$L_{\text{Po}} = \log \left\{ \prod_{i=1}^I \prod_{j=1}^J f_{\text{Po}}(x_{ij}|\theta_{ij}) \right\} \quad (3.28)$$

$$= \sum_{i=1}^I \sum_{j=1}^J \log (f_{\text{Po}}(x_{ij}|\theta_{ij})) \quad (3.29)$$

$$= \sum_{i=1}^I \sum_{j=1}^J \log \left(\frac{\theta_{ij}^{x_{ij}} e^{-\theta_{ij}}}{x_{ij}!} \right) \quad (3.30)$$

$$= \sum_{i=1}^I \sum_{j=1}^J \left[- \left\{ x_{ij} \log \frac{x_{ij}}{\theta_{ij}} - (x_{ij} - \theta_{ij}) \right\} + x_{ij} \log x_{ij} - \log x_{ij}! - x_{ij} \right] \quad (3.31)$$

$$= -\mathcal{D}_{\text{KL}}(\mathbf{X}|\Theta) + \beta_{\text{Po}} \quad (3.32)$$

ここで， β_{Po} は θ_{ij} に依存しない定数である．よって， L_{Po} を最大化することは一般化 Kullback-Leibler ダイバージェンスを最小化することに等しい．

最後に， \mathbf{X} の要素 x_{ij} が互いに独立で指数分布 $\text{Exp}(1/\theta_{ij})$ に従うことを仮定する．このとき，確率密度関数を $f_{\text{Exp}}(x_{ij}|\theta_{ij})$ とすれば，対数尤度 L_{Exp} は次のように変形することができる．

$$L_{\text{Exp}} = \log \left\{ \prod_{i=1}^I \prod_{j=1}^J f_{\text{Exp}}(x_{ij}|\theta_{ij}) \right\} \quad (3.33)$$

$$= \sum_{i=1}^I \sum_{j=1}^J \log (f_{\text{Exp}}(x_{ij}|\theta_{ij})) \quad (3.34)$$

$$= \sum_{i=1}^I \sum_{j=1}^J \log \left(\frac{1}{\theta_{ij}} e^{-\frac{x_{ij}}{\theta_{ij}}} \right) \quad (3.35)$$

$$= \sum_{i=1}^I \sum_{j=1}^J \left[- \left\{ \frac{x_{ij}}{\theta_{ij}} - \log \frac{x_{ij}}{\theta_{ij}} - 1 \right\} - \log x_{ij} - 1 \right] \quad (3.36)$$

$$= -\mathcal{D}_{\text{IS}}(\mathbf{X}|\Theta) + \beta_{\text{Exp}} \quad (3.37)$$

ここで， β_{Exp} は θ_{ij} に依存しない定数である．よって， L_{Exp} を最大化することは板倉斎藤擬距離を最小化することに等しい．

3.5 本章のまとめ

本章では，データセット DadaGP について述べた．また，ギターリフが繰り返される旋律であることを認識するための技術として，コサイン類似度及び NMF について説明した．次章では，ギタータブ譜を行列に変換する過程について述べる．そして，コサイン類似度を用いて旋律に対する類似度を定義し，提案手法である MRME 法，CRME 法，及び NRME 法について述べる．

第 4 章

提案手法

4.1 まえがき

本論文の主題であるギタータブ譜中の繰り返しパターン（ギターリフ）の抽出を実現するためには、まず DadaGP のデータセットに含まれるギタータブ譜を何らかの形で数値に変換する必要がある。そのため、4.2 節ではギタータブ譜に含まれる音価と押弦の位置という 2 つの情報のみを非負行列に変換する過程について述べる。その後、ギターリフが繰り返されていることを認識するために、コサイン類似度を用いて、行列の類似度を表す類似度行列について述べる。そして、4.3 節及び 4.4 節では、類似度行列を用いてギターリフの開始位置及び長さを求める MRME 法、CRME 法について述べる。4.5 節では、ギタータブ譜中で最も繰り返される旋律を求める NRME 法について述べる。

4.2 ギタータブ譜の数値化

本論文では以降、実験を含め 4/4 拍子の楽曲のギタータブ譜のみを解析の対象に限定する。しかしながら、提案手法は他の 5/4 拍子や 6/8 拍子等にも一般性を失うことなく同様の方法で適用可能である。

4.2.1 ギタータブ譜の最小音価に基づく圧縮

ギタータブ譜の小節数を $N \in \mathbb{N}$ で表し、小節番号を指定するためのインデクスとして n ($n = 1, 2, \dots, N$) を用いる。ギターの弦の数を $S \in \mathbb{N}$ で表し、第何弦かを指定するためのインデクスとして s ($s = 1, 2, \dots, S$) を用いる。ギターの最大のフレット番号を $F \in \{0\} \cup \mathbb{N}$ で表し、フレット番号を指定するためのインデクスとして f ($f = 0, 1, \dots, F$) を用いる。

3.2.2 項で述べたように DadaGP 形式で表されたギタータブ譜の各音符・休符の音価が wait によって指定されている。これを利用してタブ譜上に新たな時間の概念を導入する。

あるギタータブ譜の中で、最も短い音価が l tick ($1 \leq l \leq 3840$) であったとする。この時、

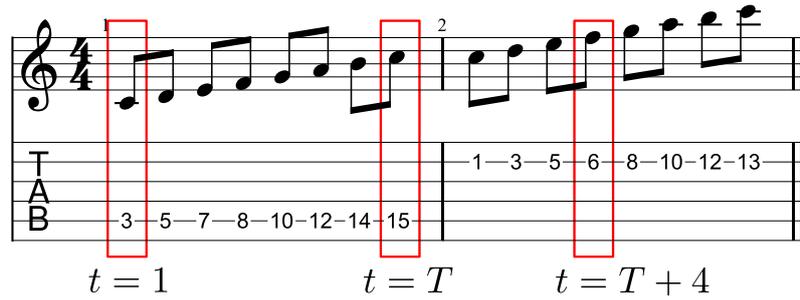


Fig. 4.1. Example of new time index.

圧縮後のギタータブ譜における1小節の長さ (tick 数) を T とおき、次式で定義する。

$$T = \text{gcd}(3840, l) \quad (4.1)$$

ここで $\text{gcd}(\cdot, \cdot)$ は入力された2つの自然数の最大公約数を表す。例えば、Fig. 2.8において最も短い音符は8分音符で、その長さは480 tickである。よって、Fig. 2.8の場合は $T = 8$ となる。

これ以降、ギタータブ譜の1小節の長さを T tick と記述する。そして、その時間を指定するためのインデクスとして t を用いる。この時間の概念は、次のように定義される。

- 1小節の長さを T とする。ただし T は式(4.1)によって定まるものである。
- 1小節目における1拍目を $t = 1$ とする。

例えば、Fig. 4.1において $t = 1$ は第5弦・第3フレットを指し、 $t = T + 4$ は第2弦・第6フレットを指す。この定義より、ギタータブ譜の終端を表す時刻は $t = NT$ となるが、この時刻を便宜上 E と定義する。また、 n 小節目の1拍目を表すための時間インデクス τ_n を次式で定義する。

$$\tau_n = 1 + (n - 1)T \quad (4.2)$$

ここで、新たに時間の概念を導入した理由について述べる。次項ではギタータブ譜を行列に変換する。その際に、新たな時間を用いることで、音価と押弦の位置の情報を失なうことなく、最もサイズの小さい行列を構成することができるからである。ギタータブ譜上で、ある弦のあるフレットの状態は、音の長さ、音の開始位置の2つの情報によって表すことができる。この運指の変換例を Fig. 4.2 に示す。ここで、 Q_{ex} の列インデクスを時刻として扱う。音の長さを、 Q_{ex} の1行目における成分1が連続する数によって表し、音の開始位置を、 Q_{ex} の2行目の成分1によって表す。DadaGP の tick 数に従った場合、Fig. 4.3 に示すように、 \hat{Q}_{ex} は1小節を表すだけで3840列を必要とする。これは冗長であるため、新たな時間を導入し、ギタータブ譜を表す行列を圧縮する。

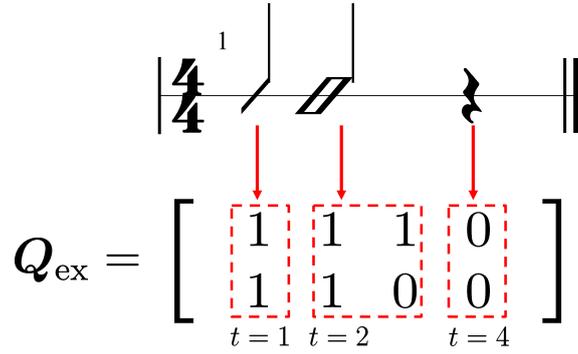


Fig. 4.2. Conversion of guitar tablature into matrix in units of new time index.

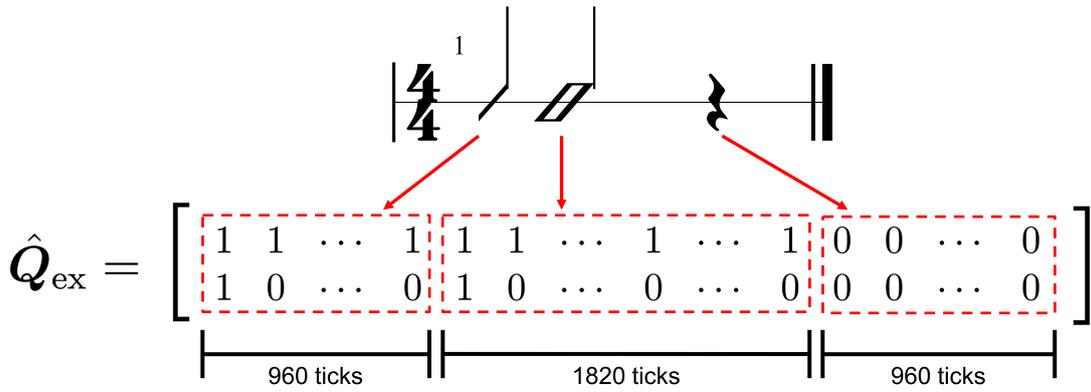


Fig. 4.3. Conversion of guitar tablature into matrix in units of DadaGP format.

4.2.2 ギタータブ譜の行列化

本項では、4.2.1 項で述べた新たな時間インデックスを用いて、ギタータブ譜をベクトル又は行列に変換する過程について述べる。まず、時刻 t における第 s 弦・第 f フレットの状態 c_{stf} を次のように定義する。

$$c_{stf} = \begin{cases} 1 & (\text{時刻 } t \text{ に第 } s \text{ 弦・第 } f \text{ フレットが押下されている}) \\ 0 & \text{otherwise} \end{cases} \quad (4.3)$$

ただし、ギタータブ譜上で $f = 0$ (第0フレット) の表記があれば、(一般に第0フレットは開放弦と呼ばれ、押下せずに弦を弾くことであるが) その場合も $c_{st0} = 1$ とする。次に、時刻 t における第 s 弦の全フレットの押下状態を表すベクトル $\mathbf{c}_{st} \in \{0, 1\}^{F+1}$ を次のように定義する。

$$\mathbf{c}_{st} = [c_{st0}, c_{st1}, \dots, c_{stF}]^T \quad (4.4)$$

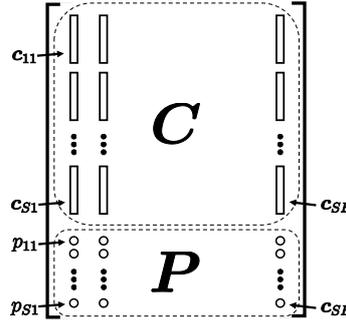


Fig. 4.4. Tablature matrix.

次に、全時刻の全弦と全フレットの押下状態を表す押下行列 $C \in \{0, 1\}^{S(F+1) \times E}$ を次のように定義する。

$$C = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1E} \\ c_{21} & c_{22} & \cdots & c_{2E} \\ \vdots & \vdots & \ddots & \vdots \\ c_{S1} & c_{S2} & \cdots & c_{SE} \end{bmatrix} \quad (4.5)$$

次に、時刻 t において s 弦を弾いた瞬間の時刻（ピッキングした瞬間）とそうではない時刻（押下を続けて音を持続させている状態）を区別するための変数 p_{st} 及び p_{st} を次のように定義する。

$$p_{st} = \begin{cases} 1 & \text{(時刻 } t \text{ の瞬間に第 } s \text{ 弦をピッキングしている)} \\ 0 & \text{otherwise} \end{cases} \quad (4.6)$$

さらに、この変数を並べたピッキング行列 $P \in \{0, 1\}^{S \times E}$ を次のように定義する。

$$P = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1E} \\ p_{21} & p_{22} & \cdots & p_{2E} \\ \vdots & \vdots & \ddots & \vdots \\ p_{S1} & p_{S2} & \cdots & p_{SE} \end{bmatrix} \quad (4.7)$$

このようにして構築した押下行列 C 及びピッキング行列 P を行方向に結合することで、Fig. 4.4 に示すようなギタータブ譜に含まれる音価と押弦の位置の情報を全て含むタブ譜行列 $M \in \{0, 1\}^{S(F+2) \times E}$ が次式のように定義できる。

$$M = \begin{bmatrix} C \\ P \end{bmatrix} \quad (4.8)$$

続いて、タブ譜行列の補数（論理反転）のような情報を持つ休符行列を定義する。まず、 $\gamma \in \mathbb{N}$ とし、ベクトル $\mathbf{v} \in \mathbb{R}^\gamma$ を引数とするデルタ関数 $\delta_\gamma(\mathbf{v}) \in \{0, 1\}$ を次のように定義する。

$$\delta_\gamma(\mathbf{v}) = \begin{cases} 1 & (\mathbf{v} = \mathbf{0}) \\ 0 & \text{otherwise} \end{cases} \quad (4.9)$$

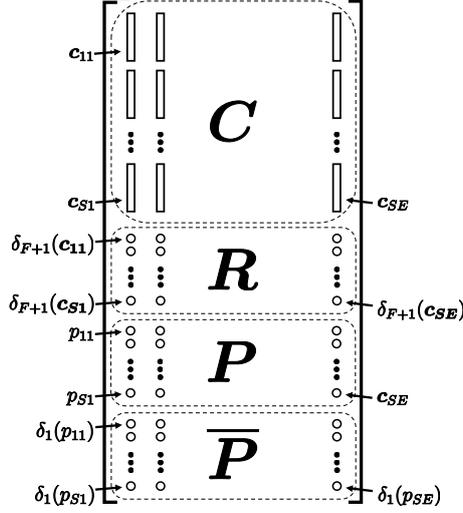


Fig. 4.5. Standard tablature matrix.

次に，弦が押さえられていないことを示す休符行列 $\mathbf{R} \in \{0, 1\}^{S \times E}$ 及びピッキング行列の補数行列 $\bar{\mathbf{P}} \in \{0, 1\}^{S \times E}$ を次のように定義する．

$$\mathbf{R} = \begin{bmatrix} \delta_{F+1}(\mathbf{c}_{11}) & \delta_{F+1}(\mathbf{c}_{12}) & \cdots & \delta_{F+1}(\mathbf{c}_{1E}) \\ \delta_{F+1}(\mathbf{c}_{21}) & \delta_{F+1}(\mathbf{c}_{22}) & \cdots & \delta_{F+1}(\mathbf{c}_{2E}) \\ \vdots & \vdots & \ddots & \vdots \\ \delta_{F+1}(\mathbf{c}_{S1}) & \delta_{F+1}(\mathbf{c}_{S2}) & \cdots & \delta_{F+1}(\mathbf{c}_{SE}) \end{bmatrix} \quad (4.10)$$

$$\bar{\mathbf{P}} = \begin{bmatrix} \delta_1(p_{11}) & \delta_1(p_{22}) & \cdots & \delta_1(p_{1E}) \\ \delta_1(p_{21}) & \delta_1(p_{21}) & \cdots & \delta_1(p_{2E}) \\ \vdots & \vdots & \ddots & \vdots \\ \delta_1(p_{S1}) & \delta_1(p_{S1}) & \cdots & \delta_1(p_{SE}) \end{bmatrix} \quad (4.11)$$

最後に，押下行列 \mathbf{C} ，休符行列 \mathbf{R} ，ピッキング行列 \mathbf{P} ，及びピッキング行列の補数行列 $\bar{\mathbf{P}}$ を結合した標準タブ譜行列 $\mathbf{M}_{\text{std}} \in \{0, 1\}^{S(F+4) \times E}$ を次のように定義する．

$$\mathbf{M}_{\text{std}} = \begin{bmatrix} \mathbf{C}^T & \mathbf{R}^T & \mathbf{P}^T & \bar{\mathbf{P}}^T \end{bmatrix}^T \quad (4.12)$$

この行列の概略図を Fig. 4.5 に示す．ここで，任意の s 及び t に関して，ベクトル $[\mathbf{c}_{st}^T, \delta_{F+1}(\mathbf{c}_{st}), p_{st}, \delta_1(p_{st})]^T$ の全ての成分の総和は 2 となる．よって，標準タブ譜行列 \mathbf{M}_{std} における任意の列ベクトルの成分の総和は必ず $2S$ となる．

ここで，押下行列 \mathbf{C} の各成分を 0 と 1 のみで構成した理由について述べる．フレット数 f は 0 以上 F 以下の値を取りうることから，時刻 t ，第 s 弦の疑似フレットの状態 \tilde{c}_{st} を次のように定義することもできる．

$$\tilde{c}_{st} = \begin{cases} f + 1 & (\text{時刻 } t \text{ に第 } s \text{ 弦のあるフレットが押下されている}) \\ 0 & \text{otherwise} \end{cases} \quad (4.13)$$

このようにフレットの状態を定義すれば、Figs. 2.2 の第 1 小節目における擬似押下行列 \tilde{C} は次のように表すことができる。

$$\tilde{C} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 3 & 3 & 3 & 3 & 3 & 3 & 5 & 6 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (4.14)$$

しかし、この定義では不都合が生じる。次項では押下行列をベクトルに変換し、押下行列に対してコサイン類似度を求める。この時、擬似押下行列を用いれば、3.3 節で述べた \cos 類似度の特徴的な性質によって、類似度が特定の成分に依存する場合が生じる。この問題を解決するために押下行列 C の各成分を式 (4.5) に示すように 0 及び 1 のみで構成した。

4.2.3 タブ譜行列のベクトル化

本項では、タブ譜行列の一部を抽出する関数及びベクトルに変換する関数について述べる。タブ譜行列 $M \in \{0, 1\}^{S(F+2) \times E}$ について、列ベクトルを $M = (\mathbf{m}_1 \ \mathbf{m}_2 \ \cdots \ \mathbf{m}_E)$ のように \mathbf{m}_t で定義する。ある $L \in \mathbb{N}$ をギターリフの小節数とする時、 M の t 列目から $t + LT - 1$ 列の列ベクトル LT 個を用いて新たに行列を構成する関数 $\text{mat}(M, t, L)$ を次のように定義する。

$$\text{mat}(M, t, L) = (\mathbf{m}_t \ \mathbf{m}_{t+1} \ \cdots \ \mathbf{m}_{t+LT-1}) \quad (4.15)$$

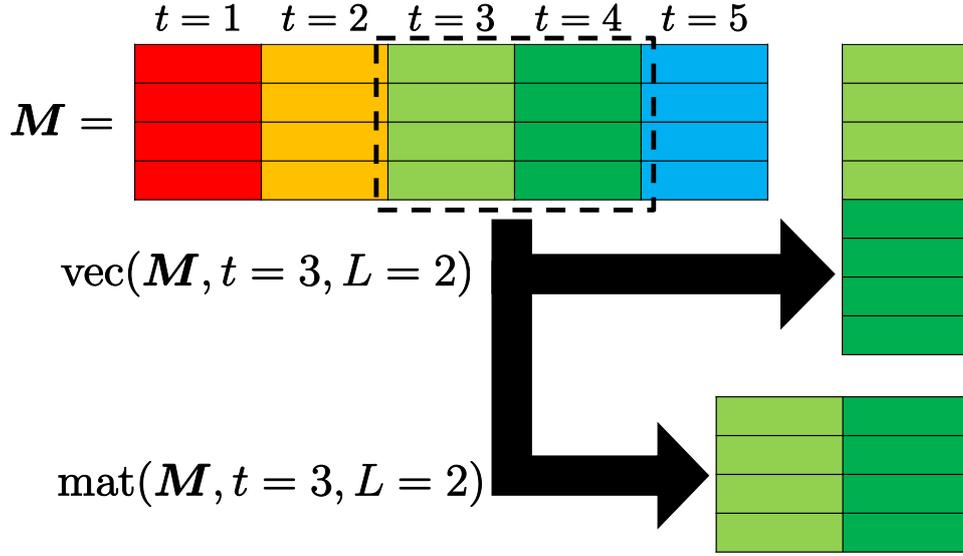
さらに、関数 $\text{mat}(M, t, L)$ で構成される行列をベクトル化する関数 $\text{vec}(M, t, L)$ を次のように定義する。

$$\text{vec}(M, t, L) = (\mathbf{m}_t^T \ \mathbf{m}_{t+1}^T \ \cdots \ \mathbf{m}_{t+LT-1}^T)^T \quad (4.16)$$

式 (4.15) 及び (4.16) の変換の概略図を Fig. 4.6 に示す。 M の t 列目は時刻 t におけるフレット及びピッキングに関する情報を持つ。よって、 $\text{mat}(M, t, L)$ は時刻 t から長さ L 小節のフレット及びピッキングに関する情報を保持した行列となる。これと同様に、 $\text{vec}(M, t, L)$ は時刻 t から L 小節の長さのフレット及びピッキングに関する情報を保持したベクトルとなる。同様の手順で押下行列 C 及びピッキング行列 P に対しても式 (4.15) 及び式 (4.16) を適用することができる。

4.2.4 タブ譜行列の類似度行列

後述の提案手法である MRME 法及び CRME 法では、式 (3.2) で定義したコサイン類似度を用いて複数小節間の類似度を測る。この際に必要となる $\sigma_{t_1, t_2}^{(M, L)} \in \mathbb{R}$ を (t_1, t_2) 成分にもつ


 Fig. 4.6. Operations of $\text{vec}(M, t, L)$ and $\text{mat}(M, t, L)$ when $t = 3$, $T = 1$, and $L = 2$.

類似度行列 $\Sigma^{(M,L)} \in \mathbb{R}^{(E-LT+1) \times (E-LT+1)}$ を次式で定義する.

$$\Sigma^{(M,L)} = \begin{bmatrix} \sigma_{1,1}^{(M,L)} & \sigma_{1,2}^{(M,L)} & \cdots & \sigma_{1,(E-LT+1)}^{(M,L)} \\ \sigma_{2,1}^{(M,L)} & \sigma_{2,2}^{(M,L)} & \cdots & \sigma_{2,(E-LT+1)}^{(M,L)} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{(E-LT+1),1}^{(M,L)} & \sigma_{(E-LT+1),2}^{(M,L)} & \cdots & \sigma_{(E-LT+1),(E-LT+1)}^{(M,L)} \end{bmatrix} \quad (4.17)$$

$$\sigma_{t_1,t_2}^{(M,L)} = \mathcal{S}_{\cos}(\text{vec}(M, t_1, L), \text{vec}(M, t_2, L)) \quad (4.18)$$

ここで, $t_1 = 1, 2, \dots, E - LT + 1$ 及び $t_2 = 1, 2, \dots, E - LT + 1$ である. $\sigma_{t_1,t_2}^{(M,L)}$ は, 時刻 $t = t_1$ から始まる長さ L 小節の旋律及び時刻 $t = t_2$ から始まる長さ L 小節の旋律がコサイン類似度を用いて, どの程度に類似しているかを表している. ここで, $\Sigma^{(M,L)}$ の t_1 行目に $\sigma_{ij}^{(M,L)} = 1$ を満たす t_2 ($t_1 \neq t_2$) が存在したとする. これは, 時刻 t_1 から始まる長さ L 小節の旋律が異なる時刻にも存在することを表している. 同様に, ピッキング行列についても類似度行列 $\Sigma^{(P,L)}$ を求めることができる. ここで, Fig. 2.10 の旋律 8 小節によって構成されるタブ譜行列の類似度行列 $\Sigma^{(M,L)}$ 及びピッキング行列の類似度行列 $\Sigma^{(M,L)}$ を Fig. 4.7 の (a) に示す. タブ譜行列における類似度行列の 33 列目 (時刻では τ_3) に注目すれば, 33, 65, \dots , 257 行目において, 高い類似度が現れていることが分かる. これによって, $t = 33 = \tau_3$ から始まる 1 小節の旋律が何度も繰り返されていることが分かる. Fig. 4.7 の (b) に示すピッキング行列に関する類似度行列についても同様のことがいえる.

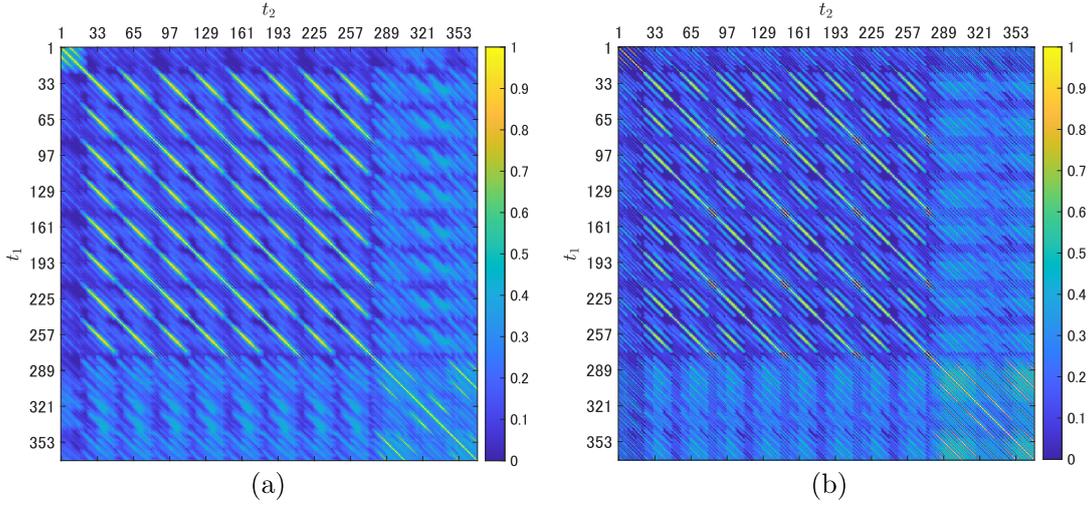


Fig. 4.7. Example of similarity matrix for “Back in Black” by AC/DC and $L = 1$: (a) $\Sigma^{(M,L)}$ and (b) $\Sigma^{(P,L)}$.

4.3 MRME 法

本節では、MRME 法について説明する。この手法では、ギタータブ譜中のギターリフは次の仮定を満たす繰り返しパターンとして定義し、ギターリフの抽出を行う。

- ギターリフの長さは $2T$, $4T$, 及び $8T$ (すなわち, 2 小節, 4 小節, 及び 8 小節) のいずれかである
- ギターリフの開始位置は $t = 1 + (n - 1)T = \tau_n$ (すなわち, 各小節の先頭) で表される
- ギターリフはそのギタータブ譜の中でバックングを除き, ギタータブ譜の冒頭から見て最初に繰り返される類似した旋律である

一般的にギターリフの長さが $6T$ で表されるものや, lT ($l \in \mathbb{N}$) の形で書き表されないものも存在する。しかし, その長さのギターリフは稀であるため, MRME 法では抽出する対象に含めない。

MRME 法では, Fig. 4.8 に示すフローチャートに従ってギターリフ抽出を行う。まず, Fig. 4.8 の Section A について説明する。ここでは正しいギターリフの開始位置を判別する。判定内容は次の命題に示す通りである。

$$\exists n' \in \mathbb{N}, \sigma_{\tau_n, \tau_{n'}}^{(\mathbf{P}, 1)} \geq \lambda \wedge n < n' \leq N \quad (4.19)$$

ここで, n' は類似する旋律が開始される小節数となる。また, λ はギターリフの開始位置であるピッキング行列 \mathbf{P} が, どの程度類似している必要があるか, を表すパラメータ ($0 \leq \lambda \leq 1$) である。式 (4.19) において, 命題 $\sigma_{\tau_n, \tau_{n'}}^{(\mathbf{P}, 1)} \geq \lambda$ が真であるとは, 時刻 τ_n から始まる長さ 1 小節のピッキング行列及び時刻 $\tau_{n'}$ から始まる長さ 1 小節のピッキング行列がコサイン類似度に

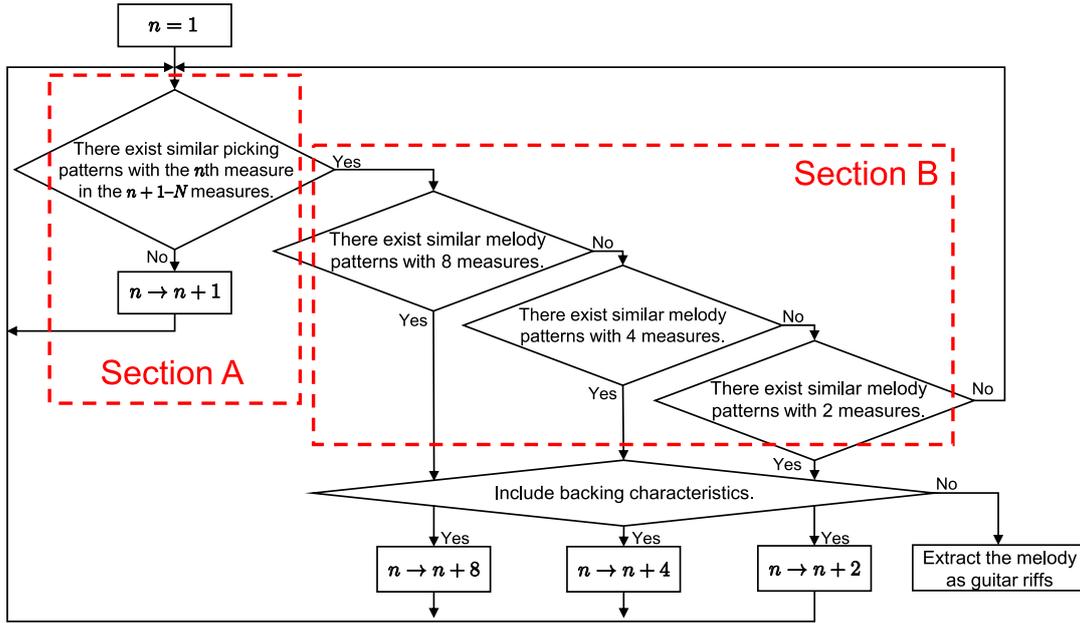


Fig. 4.8. Flow chart of MRME method.

よって λ 以上の類似度であることを表す. $n = 1$ を初期値とし, 命題が真の場合はギターリフが開始される小節位置が得られたとして Section B に進む. 命題が偽の場合は $n + 1$ について再び判定を行う.

例として, AC/DC の “Back in Black” (Fig. 2.10) について, 1 小節目がギターリフの開始位置であるかどうか判断する手順について述べる. MRME 法はまず, Fig.4.7 の (b) における $\sigma_{\tau_1, \tau_{n'}} (n' = 2, 3, \dots)$ に注目する. このギタータブ譜においては $T = 16$ であることから, $\tau_{n'} = 17, 33, \dots (n' = 2, 3, \dots)$ である. これを Fig. 4.9 に示す. 実際に計算を行えば, $\sigma_{\tau_1, \tau_2} = \frac{1}{\sqrt{2}}$ が最大値となる. よって, $\lambda > \frac{1}{\sqrt{2}}$ を満たす λ を定めれば, このセクションは 1 小節目がギターリフの開始位置でないと判断する. $\lambda \leq \frac{1}{\sqrt{2}}$ を満たす λ を定めれば, このセクションは 1 小節目がギターリフの開始位置であるとして, Section B に進む.

Section B では, ギターリフの長さ $L (L = 8, 4, 2)$ を判定する. Section A で正しいギターリフの開始位置である n 小節目 (時刻では τ_n) が得られたとする. この時, 長さ L 小節の旋律が繰り返されているかは次の命題によって判定される.

$$\exists n' \in \mathbb{N}, \sigma_{\tau_n, \tau_{n'}}^{(C, L)} \geq \mu \wedge n' \equiv n \pmod{L} \wedge n < n' \leq N - L + 1 \quad (4.20)$$

n' は類似する旋律が開始される小節数となる. また, μ は類似する旋律の開始位置である押下行列 C が, どの程度類似している必要があるか, を表すパラメータ ($0 \leq \mu \leq 1$) である.

式 (4.20) において, $\sigma_{\tau_n, \tau_{n'}}^{(C, L)} \geq \mu$ が真であるとは, 時刻 τ_n から始まる長さ L 小節の押下行列及び時刻 $\tau_{n'}$ から始まる長さ L 小節の押下行列がコサイン類似度によって μ 以上の類似度であることを表す. また, $n' \equiv n \pmod{L} \wedge n < n' \leq N - L$ が真であるとは, $n' = n + L, n + 2L, n + 3L, \dots$ かつ $n < n' \leq N - L + 1$ の範囲を動くことを表す. この命

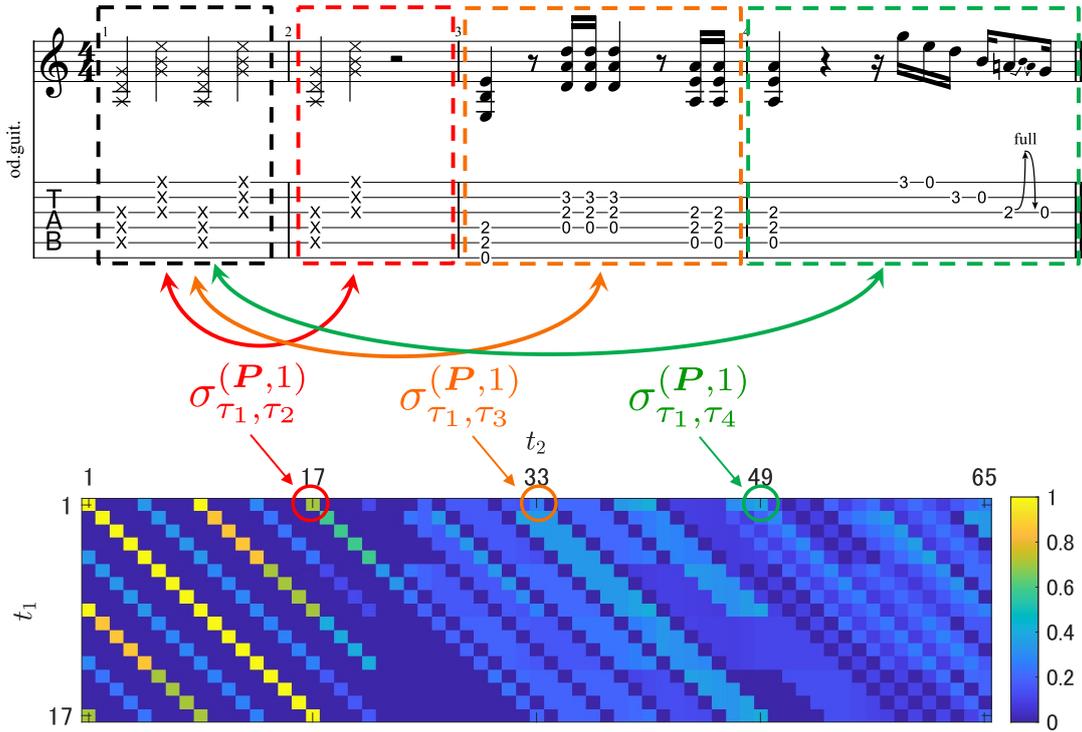


Fig. 4.9. Location of the component in similarity matrix.

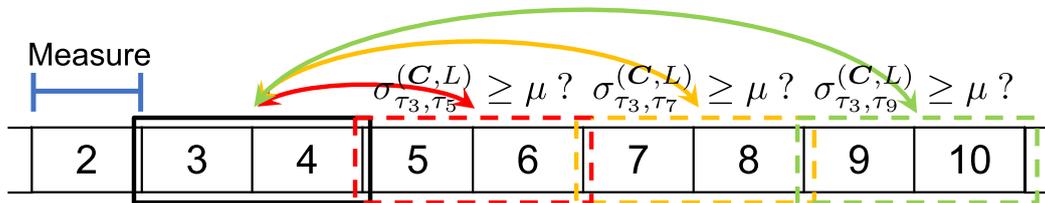


Fig. 4.10. Overview of Section B in MRME method when $n = 3$ and $L = 2$.

題において、 $n = 3, L = 2$ における概略図を Fig. 4.10 に示す．ここで、ブロックで囲われた数字は何小節目であるかを表すものである．ここで、 $n' = n + L + 1$ のような $n' \equiv n \pmod L$ が真でない n' については判定を行わない．これは、リフの長さ L が $L = 8, 4, 2$ であるという仮定に基づくもので、これ以外の想定されていないメロディの抽出を防ぐためである．

Section B では、 $L = 8, 4, 2$ の順番で命題の判定を行う．これは、Fig. 2.5 における (D) のリフを抽出する場合、 $L = 2, 4, 8$ の順に命題判定を行えば、Section B は AB を出力し、正しいギターリフが抽出できないからである．経験的に 8 小節より長い旋律は短い旋律に分割できる．よって $L = 16$ の場合は探索を行わない．

最後に、Section B で抽出された複数の時間で類似する旋律パターンはバックギタリフかどうかの判定を行う．ここで、時刻 τ_n から長さ L 小節の旋律を抽出したとする．このとき、バック

ングであるかどうかの判定は次の不等式によって行う。

$$\text{Rank}(\text{mat}(\mathbf{C}, \tau_n, L)) \geq \xi \quad (4.21)$$

ここで、 $\xi \in \mathbb{Z}_{\geq 0}$ はバッキングであると判断するためのパラメータである。例として、 $\xi = 2$ と定めた場合、Fig. 2.13, Fig. 2.14, 及び Fig. 2.16 に示す 2 小節の旋律は、バッキングであると判断される。さらに、 $\xi = 3$ と定めた場合には、Fig. 2.15 に示す 2 小節の旋律もバッキングであると判断することができる。この不等式を満たさない場合はバッキングであると判断され、 $n \rightarrow n + 1$ とし、再び Section A を開始する。

4.4 CRME 法

4.3 節で述べた MRME 法はギターリフの長さを $2T$, $4T$, 及び $8T$ であるという仮定の下、ギターリフの抽出を行う手法である。しかしながら、実際のギターリフにおいては、ギターリフの長さが $6T$ で表されるものや、 lT ($l \in \mathbb{N}$) の形で書き表されないものも存在する。そこで、本節では MRME 法とは異なるギターリフの仮定を用いる CRME 法を新たに提案する。具体的には、CRME 法はギターリフの仮定を次のように定めて抽出を行う。

- ギターリフの小節数 L が分かっているものとする
- ギターリフの開始位置は $t = 1 + (n - 1)T = \tau_n$ (すなわち、各小節の先頭) である
- ギターリフはそのギタータブ譜の中でバッキングを除き、ギタータブ譜の冒頭から見て最初に繰り返される類似した旋律である

ギターリフの小節数 L が与えられた時、CRME 法は $\Sigma^{(M,L)}$ を用いてギターリフ抽出を行う。MRME 法ではピッキング行列 \mathbf{P} 及び押下行列 \mathbf{C} を用いたが、この手法が用いるのはタブ譜行列 \mathbf{M} であることに注意する。タブ譜行列 \mathbf{M} を用いることで、MRME 法のようなアルゴリズムを用いることなく、単純な演算によってギターリフの抽出を可能としている。

まず、次の集合 U 及び V を構成する。

$$U = \left\{ n \in \mathbb{N} \mid \exists t \in \mathbb{N}, \sigma_{\tau_n, t}^{(M,L)} \geq \nu \wedge \tau_n < t \leq E - LT + 1 \right\} \quad (4.22)$$

$$V = \{ m \in \mathbb{N} \mid m \in U \wedge \text{Rank}(\text{mat}(\mathbf{C}, \tau_m, L)) \geq \xi \} \quad (4.23)$$

ここで、 ν は類似する旋律におけるタブ譜行列 \mathbf{M} が、どの程度類似している必要があるか、を表すパラメータ ($0 \leq \nu \leq 1$) である。集合 U は、小節数 L の類似旋律が存在する旋律の開始位置の集合である。ただし、この集合は式 (4.20) とは異なり、mod による制約が無い。これによる違いは、Fig. 4.11 に示すように、全ての時刻における旋律を探索することである。これは、Fig. 2.5 における (G) のギターリフを抽出するためであり、MRME 法にはない特徴である。

式 (4.23) において、 $\xi \in \mathbb{Z}_{\geq 0}$ はバッキングであると判断するためのパラメータであり、集合 V は集合 U によって定まる旋律のうち、バッキングを除いた集合である。CRME 法は、集合

Fig. 4.11. Overview of CRME method when $n = 1$ and $L = 2$.

V の最小元をギターリフの開始位置とする．小節数 L が既に与えられていることから，これによって1つの旋律が定まる．ただし，集合 V が空である場合はギターリフが存在しないものとする．

4.5 NRME 法

これまでに述べた MRME 法及び CRME 法はギタータブ譜において，初めに繰り返される旋律をギターリフとして抽出する手法である．しかし，ギターリフは初めに繰り返される旋律とは限らない．そこで，非負行列に含まれるパターンを低ランク近似で抽出することができる NMF に注目する．ギターリフが何度も繰り返されるという特徴を持つため，NMF を用いることで，初めに繰り返されないギターリフも抽出を行うことができると考えられる．そこで本節では，NMF を用いてギターリフ抽出を行う NRME 法について述べる．NRME 法では，次のような仮定をギターリフに求める．

- ギターリフの小節数 L は分かっているものとする
- ギターリフの開始位置は $t = 1 + (n - 1)T = \tau_n$ (すなわち，各小節の先頭) である
- ギターリフはそのギタータブ譜において，最も繰り返される旋律である．ただし，初めに繰り返される旋律とは限らない．

NRME 法におけるギターリフ抽出の手順について説明する．まず次の集合 X 及び行列

$\mathbf{Y} \in \{0, 1\}^{(S(F+4)LT) \times (E-LT+1)}$ を構成する.

$$X = \{n \in \mathbb{N} \mid \text{Rank}(\text{mat}(\mathbf{C}, \tau_n, L)) \geq \xi \wedge 1 \leq n \leq N - L + 1\} \quad (4.24)$$

$$\mathbf{Y} = (\mathbf{y}_1 \ \mathbf{y}_2 \ \cdots \ \mathbf{y}_t \ \cdots \ \mathbf{y}_{E-LT+1}) \quad (4.25)$$

$$\mathbf{y}_t = \text{vec}(\mathbf{M}_{\text{std}}, t, L) \quad (4.26)$$

集合 X はバッキングでないと判断された小節数 L の旋律の開始位置の集合, \mathbf{y}_t は時刻 t から始まる長さ L 小節のタブ譜の情報を持ったベクトル, 及び \mathbf{Y} はそれを並べたものである.

次に, 誤差関数 $\mathcal{D}(\cdot)$ を用いて \mathbf{Y} を $\hat{\mathbf{W}} \in \mathbb{R}^{(S(F+4)LT) \times \hat{K}}$ 及び $\hat{\mathbf{H}} \in \mathbb{R}^{\hat{K} \times (E-LT+1)}$ の積で近似する. このとき, パラメータ \hat{K} は楽曲の大部分を構成する旋律の数と解釈することができる. また, $\hat{\mathbf{W}}$ はギタータブ譜に頻出する旋律を各列ベクトルに持ち, $\hat{\mathbf{H}}$ は $\hat{\mathbf{W}}$ が持つ旋律の大きさを表すと解釈することができる. NMF は, 行列に頻出するパターンを高精度に復元する. これを認識するために, 誤差を表す行列 $\mathbf{D} = \mathbf{Y} - \hat{\mathbf{W}}\hat{\mathbf{H}} \in \mathbb{R}^{(S(F+4)LT) \times (E-LT+1)}$ を構成する.

頻出する旋律の開始時刻を \hat{t} とすれば, \mathbf{D} の第 \hat{t} 列目のベクトル $\mathbf{d}_{\hat{t}}$ の大きさは小さくなるはずである. NRME 法では X の元のうち, 最も $\|\mathbf{d}_t\|_2$ が小さくなる t をギターリフの開始位置とする. ここで $\|\cdot\|_2$ は L^2 ノルムを表す. ギターリフの長さ L が与えられていることから, 以上の手順によってギターリフ抽出が完了する.

$\hat{\mathbf{W}}$ の各列がどんなベクトルであろうとも $\hat{\mathbf{H}}$ の成分 $\hat{h}_{\hat{k}t}$ が全ての \hat{k} に対して $\hat{h}_{\hat{k}t} = 0$ を満たすならば, $\hat{\mathbf{W}}\hat{\mathbf{H}}$ の t 列目は零ベクトルとなる. これは, NMF が零ベクトルを簡単に復元できることを示している. この提案手法が標準タブ譜行列 \mathbf{M}_{std} を用いたのは \mathbf{Y} の各列ベクトルにおいて, 零ベクトルの出現を防ぐためである. 標準タブ譜行列 \mathbf{M}_{std} の任意の列ベクトルの成分の総和は $2S$ となる. よって, 任意の \mathbf{y}_t の成分の総和は $2SLT$ となり, \mathbf{Y} の各列ベクトルにおいて, 零ベクトルは出現しない.

4.6 本章のまとめ

本章では, タブ譜の情報を保持した行列である押下行列 \mathbf{C} , ピッキング行列 \mathbf{P} , タブ譜行列 \mathbf{M} , 休符行列 \mathbf{R} , ピッキング行列の補数行列 $\bar{\mathbf{P}}$, 及び標準タブ譜行列 \mathbf{M}_{std} を定義した. そして, それらを用いて, 提案手法である MRME 法, CRME 法, 及び NRME 法について述べた. 次章では, 提案手法である, MRME 法, CRME 法, 及び NRME 法を用いたギターリフ抽出実験を行い, 評価を行う.

第 5 章

ギターリフ抽出実験

5.1 まえがき

本章では、本論文の目的である、ギターリフのデータセットを作成すること、を達成するために、4章で提案した手法によってギターリフ抽出を行う。そして、各手法の性能を評価する。まず、5.2節ではギターリフ抽出を行うタブ譜に対する条件及び正しい抽出位置について述べる。5.3節、5.4節、及び5.5節では、各手法が持つパラメータを変化させ、最大でどの程度の精度が得られるかを確認する。5.6節では、各手法がさらに良い結果を得るための方法について考察する。最後に、5.7節で本章をまとめる。

5.2 実験条件

本論文では、次の条件を全て満たすギタータブ譜をギターリフ抽出実験に用いた。

- (a) DadaGP にギタータブ譜として含まれている
- (b) 有名なギターリフをまとめた文献 [11] またはサイト [12, 13] にギターリフが記載されている
- (c) 2小節目から 24小節目が 4/4 拍子で構成されている
- (d) ギタータブ譜で指定された調弦が変則チューニングでない

(b) は、ロック音楽史に残る有名なギターリフのみを対象とすることで、その楽曲に含まれる正解のギターリフが客観的にはっきりしているものを選ぶために設けた条件である。(c) において、4/4 拍子を扱う理由は、扱うギターリフの中で最も 4/4 拍子で構成された楽曲が多いからである。また、1小節目の拍子を任意とした理由は DadaGP で配布されているギタータブ譜の中に、弱起（小節の途中から始まる楽曲）の 4/4 拍子の楽曲の 1小節目が 4/4 拍子に設定されていない楽曲が存在するためである。この場合、Fig. 5.1 に示すように、冒頭に休符を挿入することで 4/4 拍子を構成した。さらに、(d) について、変則チューニングとは、各弦の第 0 フレット（開放弦）の音高の相対関係がスタンダードチューニングと異なるもののことを

Enter Sandman
Metallica
Metallica

Standard tuning
♩ = 123

(a)

(b)

Fig. 5.1. Example of first measure conversion: (a) before and (b) after conversion [7].

指す。例えば，全ての弦の開放弦の音高をスタンダードチューニングから半音や全音下げる調弦は変則チューニングとみなさないが，第6弦だけを全音下げる調弦（ドロップDチューニング）は変則チューニングとなり，実験データに含めないものとした。DadaGP形式のデータでは，これらの変則チューニングは3.2.2項で説明したトークンにおいて第-1フレットや第-2フレットを用いて表現されているため，そのようなトークンを含むギタータブ譜を除外した。

次に，ギタータブ譜に施す前処理について説明する。まず，DadaGP形式のファイルを編集し，Fig. 5.2に示すように，タブ譜内の繰り返し記号を全て展開された形に書き直した。最後に，展開されたギタータブ譜の1~24小節目のみを抽出し，25小節目以降は全て取り除いた。この処理は，2.3節で述べた通り，ギターリフが楽曲のイントロで演奏されることを仮定したものであり，実際にギターリフを正しく抽出する精度の向上につながる前処理となっている。

次に，ギターリフ抽出における正解値について説明する。ギターリフは何度も繰り返されていることから，正しい抽出位置が複数存在しており，それらは全て正解値となる。例えば，Deep Purpleの“Smoke on the Water”はFig. 2.8に示す1~4小節の旋律が6回繰り返される。よって，ギターリフの開始位置の正解値 n は $n = 1, 5, 9, \dots, 21$ 小節目となる。そして，この開始位置から4小節の長さを抽出して初めて正解となる。また，この開始位置から8小節の長さを抽出した場合も正解とみなす。また別の例として，Metallicaの“Enter Sandman” (Fig. 5.1)においては，開始位置の正解値 n を $n = 2, 3, \dots, 13$ とする。ただし，例外として1小節目から2小節の長さの旋律を抽出する場合も正解とみなす。これは，Fig. 5.1が1小節目の1拍目から始まっていないことに起因する。さらに別の例として，Led Zeppelinの“Black Dog” (Fig. 2.12)においては，開始位置の正解値 n を $n = 6, 13$ と定める。つまり，ギターリフの1拍目が存在する小節を開始位置と定義している。そして，その開始位置から6小節の長さの旋律を抽出できれば正解とみなす。

Paranoid
Black Sabbath
Black Sabbath

Standard tuning
♩ = 160

(a)

The figure displays two versions of the guitar riff for 'Paranoid' by Black Sabbath. Part (a) shows the original riff with two repeated measures, each consisting of four measures. Part (b) shows the riff after conversion, where the repeated measures have been expanded to eight measures each. The notation includes a treble clef, a 4/4 time signature, and a tempo of 160. The guitar part is written on a six-string staff with standard tuning (E, A, D, G, B, E). The bass line is indicated by fret numbers (12, 14) and bar lines. The guitar part is marked with 'H' for hammer-ons and 'B' for bends. The expanded version (b) shows the riff repeated for a total of eight measures, with the same notation as the original version (a).

(b)

Fig. 5.2. Example of expansion of repeated measures: (a) before and (b) after conversion [7].

5.3 MRME 法の最適パラメータ探索と最高性能の調査

4.3 項で説明したように, MRME 法には事前に定めるパラメータ μ , λ , 及び ξ が存在する. 本実験では, $\mu = 0.5, 0.6, \dots, 1.0$, $\lambda = 0.50, 0.51, \dots, 1.00$, 及び $\xi = 0, 1, \dots, 5$ の範囲でパラメータを動かし, ギターリフ抽出を行う. そして, 最も高い精度でギターリフが抽出できる時のパラメータ及び抽出精度を確認する.

得られた実験結果を Fig. 5.3 に示す. まず λ の変化に対する抽出性能に注目する. 任意の μ 及び ξ に対して, 正解率が最も良いときの λ は $0.9 \sim 0.95$ に集まっていることが分かる. $\lambda = 1$ ではなかった理由は, Fig. 2.5 における (F) のリフが使用したデータに多かったことに起因する. 一方 μ に関して, 正解率が最も良かったのは, $\mu = 0.7, 0.8$ であり, λ より小さい値となった. さらに ξ に関しては, $\xi = 1, 2, 3, 4, 5$ では正解率の最大値に変化しなかったが, $\xi = 6$ において, 正解率の最大値が 10% ほど下がった. これにより, 多くのギターリフの押下行列のランクは 5 であることが分かる. $\lambda = 0.91$, $\mu = 0.7, 0.8$, 及び $\xi = 3$ の時, 最大正解率 77.9% が得られた. また, Fig. 5.3 で明らかとなった MRME 法の最適パラメータを用いて, ギターリフの長さごとの正解数を Table 5.1 に示す. $L = 1$ のギターリフが正しく抽出されて

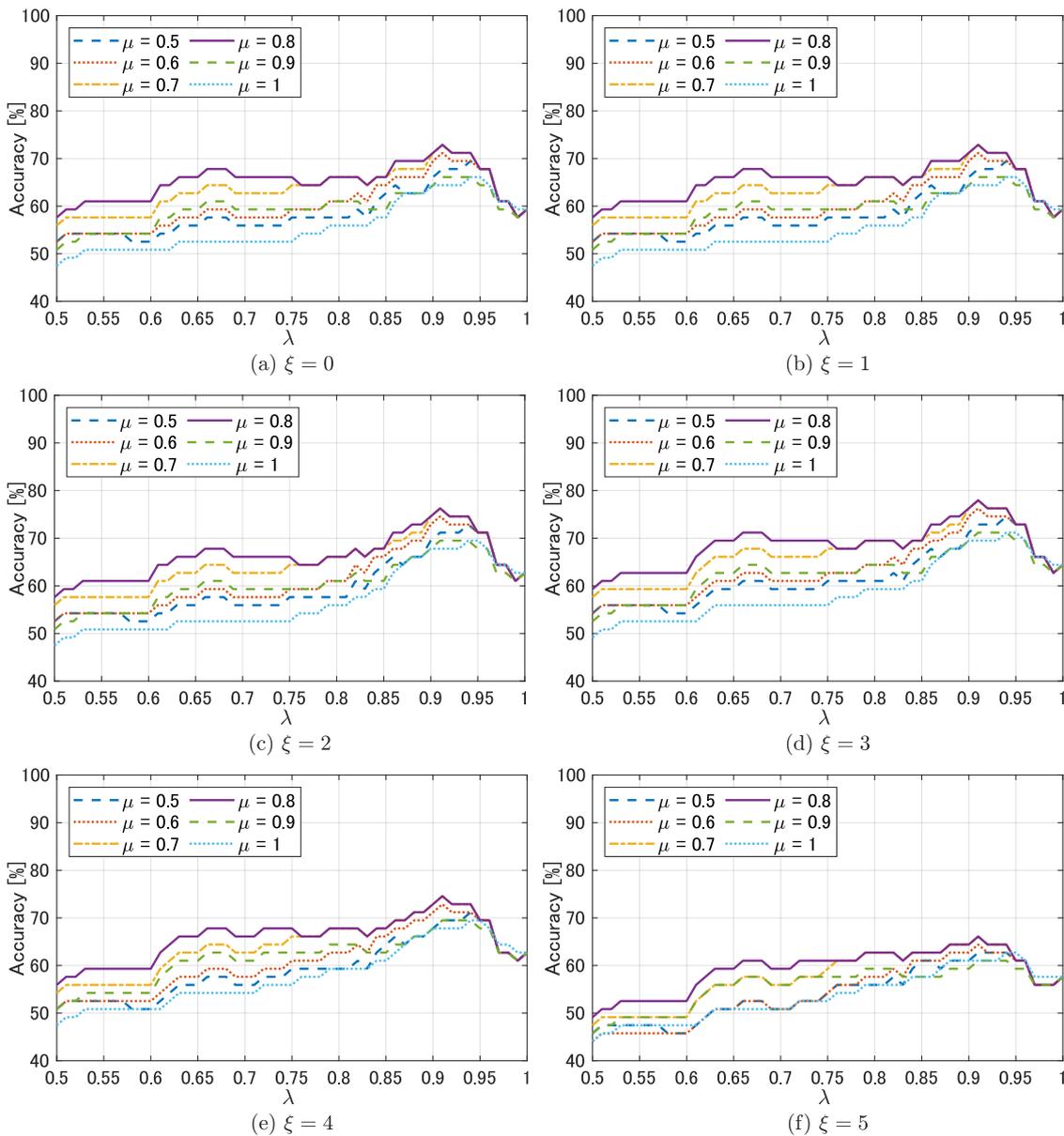


Fig. 5.3. Extraction accuracy of MRME method: (a) $\xi = 0$, (b) $\xi = 1$, (c) $\xi = 2$, (d) $\xi = 3$, (e) $\xi = 4$, and (f) $\xi = 5$.

いるのは、1小節のギターリフが繰り返され、 $L = 8, 4, 2$ のいずれかに該当したからである。ギターリフ抽出に失敗した13曲の原因は、次のように分類分けを行うことができる。

- (A) ギターリフの開始位置を正しく捉えたが、ギターリフの長さが異なる
- (B) ギターリフとは異なる旋律を抽出した
- (C) ギターリフが存在しないと判定した

(A) の例として Fig. 5.4 が挙げられる。正しいギターリフの位置は1~8小節目であったが、

Table 5.1. Number of correct results in each length of guitar riffs estimated by MRME method

Riff length	Number of tablatures	Number of correct results	Accuracy [%]
$L = 1$	3	3	100
$L = 2$	31	29	93.5
$L = 4$	19	13	68.4
$L = 6$	2	0	0.0
$L = 8$	4	1	25.0

Standard tuning
♩ = 80

Extracted melody

Fear of the Dark
Iron Maiden
Fear of the Dark

Guitar riff

Fig. 5.4. Example of incorrectly extracted results: length mismatch of guitar riffs [7].

MRME 法は 1~2 小節を抽出した。Fig. 5.4 の構造は、Fig. 2.5 の (D) における $A = C$ 及び $B = D$ の場合である。そして、この 8 小節の旋律が 1 回しか繰り返されなかったことから正しくリフを抽出できなかった。(B) の失敗の原因は、Fig. 5.5 に示すように、ギターリフが出現する前に、ギターリフではない繰り返し旋律を抽出するからである。定めたギターリフの正解位置は 9~10 小節目であるが、MRME 法は 1~2 小節目を抽出した。(C) によって失敗したのは、Fig. 2.5 における (D) のパターンである。つまり、Fig. 2.12 のタブ譜にはギターリフが存在しないと判定された。

Ace of Spades
Motorhead
Ace of Spades

Extracted melody

Standard tuning
♩ = 140

Guitar riff

Fig. 5.5. Example of incorrectly extracted results: place mismatch of guitar riffs in MRME method [7].

5.4 CRME 法の最適パラメータ探索と最高性能の調査

4.4 項で説明したように，CRME 法には事前に定めるパラメータ ν 及び ξ が存在する． $\nu = 0.50, 0.51, \dots, 1.00$ 及び $\xi = 0, 1, \dots, 5$ の範囲でパラメータを動かし，ギターリフ抽出を行った．そして，最も高い精度でギターリフが抽出できる時のパラメータ及び抽出精度を確認する．得られた実験結果を Fig. 5.6 に示す．この結果から， ν は MRME 法の λ と同様に，0.9~0.95 において，高い正解率を示した．また，MRME 法と同様に $\xi = 5$ の時，正解率の最大値が最も小さくなった． $\nu = 0.92$ 及び $\xi = 3$ の時，最大正解率 79.7% が得られた．また，このパラメータにおける，ギターリフの小節数毎の正解数を Table 5.2 に示す．

ギターリフ抽出に失敗した 12 曲の原因は，次のように分類分けを行うことができる．

- (A) ギターリフの一部を抽出しているが，開始位置が正しくない
- (B) ギターリフとは異なる旋律を抽出した
- (C) ギターリフが存在しないと判定した

MRME 法及び CRME 法の大きな違いとして，式 (4.20) 及び式 (4.22) において，mod による制約の有無がある．この違いには，Fig. 2.12 の Black Dog のギターリフのように，小節と

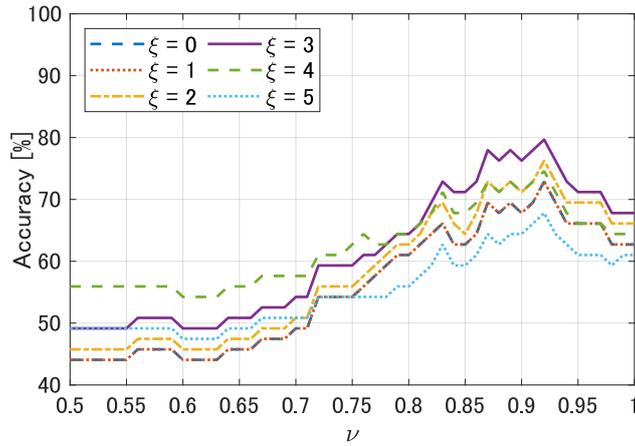


Fig. 5.6. Extraction accuracy of CRME method.

Table 5.2. Number of correct results in each length of guitar riffs estimated by CRME method.

Riff length	Number of tablatures	Number of correct results	Accuracy [%]
$L = 1$	3	2	66.7
$L = 2$	31	27	87.1
$L = 4$	19	15	78.9
$L = 6$	2	1	50.0
$L = 8$	4	2	50.0

いう区切りに捕らわれない繰り返しパターンを抽出するという目的があった。しかし、CRME法は Fig. 5.7 に示すように、4~9 小節を抽出した。この原因については 5.6 節で詳しく述べる。5.3 節で説明したパターン (B) に該当するタブ譜は、CRME 法でも正しい抽出を行うことができなかった。これは、MRME 法及び CRME 法が繰り返されている旋律の中から最も先頭にあるものを抽出していることに起因する。MRME 法の (C) に該当したタブ譜が 2 曲であったのに対し、CRME 法は 4 曲と数が多くなった。例えば、CRME 法は Fig. 5.4 のギターリフの小節数 L が $L = 8$ であることを既知としている。しかし、このギターリフは繰り返されないため、ギターリフが存在しないと判断した。これにより (C) に該当するタブ譜の数が増加した。

5.5 NRME 法の最適パラメータ探索と最高性能の調査

4.5 項で説明したように、NRME 法には事前に定める誤差関数及びパラメータ ξ 及び K が存在する。誤差関数は、式 (3.14)、式 (3.15)、及び式 (3.16) の 3 種類を用い、 $\xi = 0, 1, \dots, 5$ 及び $K = 1, 2, \dots, 5$ の範囲でパラメータを動かしてギターリフ抽出を行った。 \mathbf{W} 及び \mathbf{H} の更新回数を 20 回とし、最も高い精度でギターリフが抽出できる時のパラメータ及び抽出精度

Black Dog
Led Zeppelin
Led Zeppelin IV

Standard tuning
♩ = 150

Guitar riff

Fig. 5.7. Example of incorrectly extracted results: location mismatch of guitar riffs for in CRME method [7].

を確認する．ここで，ギターリフの長さが長いほど，旋律を並べた行列 \mathbf{Y} に出現する旋律のパターン数が多くなるはずである．つまり， L が大きくなるほど \mathbf{Y} のランクが高くなり，各 L ごとに最適な基底数 K は異なるはずである．そこで，本実験ではギターリフの長さごとに実験結果を示す．

まず，二乗 Euclid 距離を用いて得られた実験結果を Fig. 5.8 に示す．次に，一般化 Kullback–Leibler ダイバージェンスを用いて得られた実験結果を Fig. 5.9 に示す．最後に，板倉斎藤儀距離を用いて得られた実験結果を Fig. 5.10 に示す．用意した 59 曲のうち，ギターリフの長さ L が $L = 1$ であるタブ譜の数は 3 個， $L = 8$ であるタブ譜の数は 4 個である．よって， $L = 1, 8$ の場合に高い正解率が得られたとしても，他の $L = 1, 8$ のタブ譜全体に対して良いふるまいを見せる，とは言えない．タブ譜の数が最も多い $L = 2$ に注目すれば，正解率が最も高かったのは誤差関数に一般化 Kullback–Leibler ダイバージェンスを用い，パラメータが $\xi = 4$ 及び $K = 1, 2, 3, 4, 5$ の時である．この時の正解率は 66.9% であった．

誤差関数に二乗 Euclid 距離を用いた NMF は， \mathbf{Y} の各成分が正規分布から生成されることを仮定する．同様に，一般化 Kullback–Leibler ダイバージェンスはポアソン分布，板倉斎藤儀距離は指数分布から生成されることを仮定する．実際に構成した \mathbf{Y} はほとんどの成分が 0 であることから，板倉斎藤儀距離を用いた NMF が \mathbf{Y} を上手く構成し，最も良い精度でギターリフ抽出を行うことが出来ると考えていたが，結果は予想と異なるものであった．

一般化 Kullback–Leibler ダイバージェンスを用いた場合の特徴として，基底の数 K が変化してもギターリフの抽出精度が変化しないことが挙げられる．これは，他の誤差関数を用いた場合には見られない特徴である．

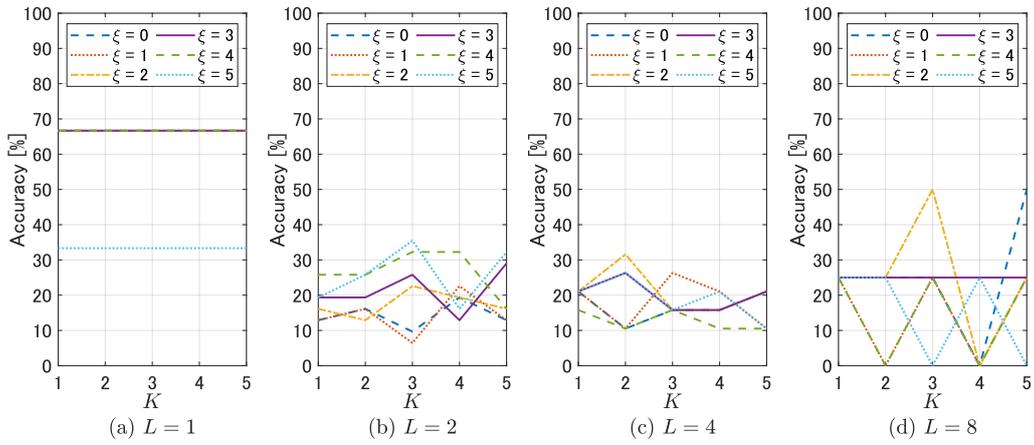


Fig. 5.8. Extraction accuracy of NRME method based on squared Euclidean distance.

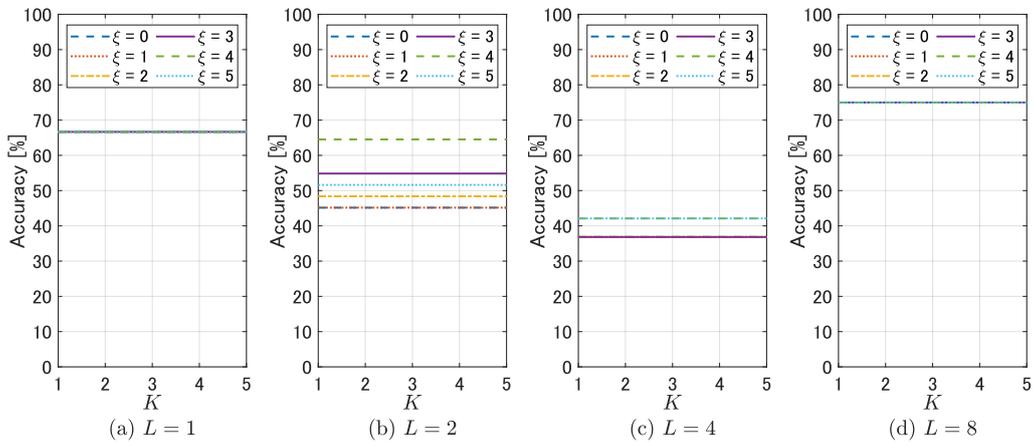


Fig. 5.9. Extraction accuracy of NRME method based on generalized Kullback–Leibler divergence.

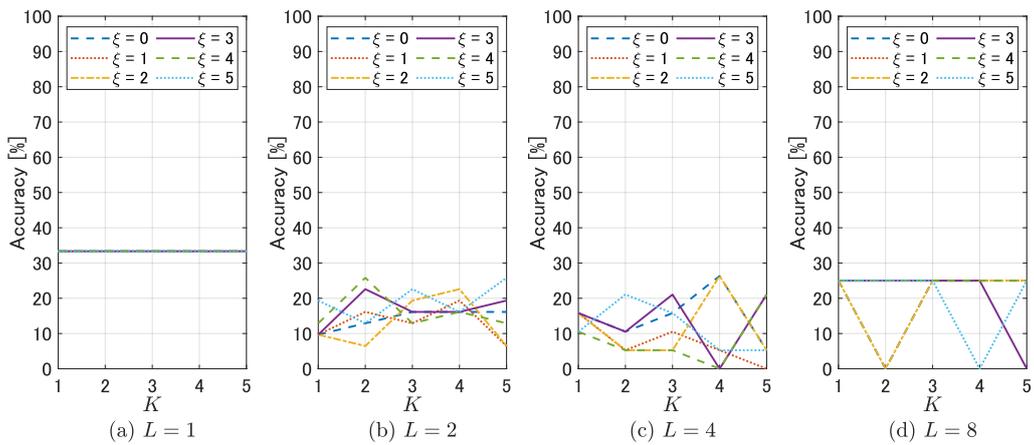


Fig. 5.10. Extraction accuracy of NRME method based on Itakura–Saito divergence.

5.6 各提案手法の考察

まず、MRME 法、CRME 法、及び CRME 法のパラメータがリフの長さに依存することについて述べる。各手法で用いた各パラメータの最適解は、リフの長さに依存するはずである。例えば、1 小節のギターリフと 8 小節のギターリフでは、1 小節のギターリフの方が類似度が高くなるはずである。さらに、1 小節のバックিংと 8 小節のバックিংでは、8 小節のバックিংの方がランクが高くなる可能性が高い。これにより、リフの長さごとにパラメータを求めることで、各手法の性能はさらに高くなることが期待される。

次に、押下行列 \mathbf{C} が一致すること及びピッキング行列 \mathbf{P} が一致することの重要度や影響度が等価ではないという可能性について述べる。MRME 法では、押下行列 \mathbf{C} 及びピッキング行列 \mathbf{P} を異なるセクションで扱った。ギターリフの長さが未知であるにもかかわらず、MRME 法の正解率が高い理由がこれにあるとするならば、CRME 法でも、押下行列 \mathbf{C} 同士的一致及びピッキング行列 \mathbf{P} 同士的一致に差を付けることで高い正解率が得られると考える。例えば、ある正の実数 k によって擬タブ譜行列 $\mathbf{M}'_k \in \{0, 1, k\}^{S(F+2) \times E}$ を次のように定める。

$$\mathbf{M}'_k = \begin{pmatrix} \mathbf{C} \\ k\mathbf{P} \end{pmatrix} \quad (5.1)$$

本論文では $k = 1$ として実験を行った。 $k \ll 1$ とすれば、類似度は押下行列 \mathbf{C} のふるまいに大きく依存し、 $k \gg 1$ とすれば類似度はピッキング行列 \mathbf{P} のふるまいに大きく依存することになる。このように擬タブ譜行列 \mathbf{M}'_k を定めることで、拡張された旋律行列を考えることができる。

さらに、CRME 法で Black Dog のギターリフ抽出に失敗した原因を考察する。原因は、コサイン類似度が休符同士の類似度を求めることができないからである。例として、Fig. 5.11 に示す 4 小節の旋律があったとする。この時、次が成り立つ。

$$\sigma_{\tau_1 \tau_3}^{(\mathbf{M}, 2)} = \sigma_{\tau_2 \tau_4}^{(\mathbf{M}, 1)} \quad (5.2)$$

この等式は、2 小節の旋律に注目していたにもかかわらず、1 小節の旋律だけに注目していることを表す。これによって、CRME 法は Led Zeppelin の “Black Dog” (Fig. 2.12) のギターリフ抽出に失敗した。この問題の解決策として、ギターリフの 1 小節目が全休符でない仮定を導入する、もしくは、休符と休符の類似度を考慮した新たな類似度を構成する方法が考えられる。

最後に、NRME 法について、ギターリフが長いほど最適な基底数 K が大きくなる、という予想とは異なる結果が得られた。結果の傾向として、Fig. 2.8 に示すような、曲の最後まで同じ旋律を繰り返すギタータブ譜においては、正しい抽出を行った。しかし、Fig. 5.5 に示すような、複数の旋律が存在する場合に正しく抽出を行うことはできなかった。これにより、 $i = 1, 2, \dots, E - LT + 1$ の範囲で \mathbf{Y} を構成するのではなく、 $i = \tau_1, \tau_2, \dots, \tau_N$ の範囲で \mathbf{Y} を構成し、出現する旋律のパターン数を減らすことで、より良い結果が得られるのではないかと考える。

Fig. 5.11. Example melody of extraction failure.

5.7 本章のまとめ

本章では、ギターリフ抽出実験を行った。ギターリフの長さを未知とする MRME 法では最大のギターリフ抽出正解率 77.9% の結果が得られ、ギターリフの長さを既知とする CRME 法はさらに高い精度である 79.7% の結果が得られた。NRME 法は、MRME 法及び CRME 法と比較して正解率が低いという結果が得られた。5.6 節では、各手法の正解率をさらに改善するための考察を行った。次章では、本論文をまとめる。

第6章

結言

本論文では、ギターリフを抽出するアルゴリズムを作成し、その精度を確認した。

1章では、本論文の動機と目的を明らかにした。本論文の動機は、ギターリフのデータセットが存在しない事である。そこで、ギターリフを抽出するアルゴリズムを作成し、その精度を確認することを本論文の目的とした。

3章では、DadaGP が提供するタブ譜のデータセットについて説明した。また、ギターリフ抽出に使用する技術であるコサイン類似度及び NMF について説明した。

4章では、提案手法で用いる押下行列、ピッキング行列、タブ譜行列、休符行列、ピッキング行列の補数行列、及び標準タブ譜行列を定義した。そして、ギターリフを抽出するための手法である、MRME 法、CRME 法、及び NRME 法について述べた。

5章では、ギターリフ抽出を行う上での条件を述べ、各手法のギターリフ抽出結果について述べた。また、各手法の正解率を改善する方法について考察した。

最後に今後の課題を述べる。各提案手法が持つパラメータがリフの長さに依存するものとして、パラメータの探索を行えば、さらに良い精度で抽出を行うことが出来ると考えられる。MRME 法が高い正解率を示した要因として、押下行列及びピッキング行列を異なるセクションで扱ったことが挙げられる。これを CRME 法に応用し、タブ譜行列を拡張した擬タブ譜行列を定め、新たなパラメータと共にギターリフ抽出を行うことでさらに高い精度が得られる可能性がある。NRME 法においては、NMF によって分解される行列を新たに定める必要がある。本論文で抽出したギターリフを用いることで、1.2 節で述べた CVAE の実装も考えられる。

謝辞

本論文は、香川高等専門学校電気情報工学科北村研究室にて行われた研究に基づくものです。

まず、本研究を進めるにあたり、ご多忙のところ熱心にご指導くださいました指導教員の北村大地講師に心より感謝申し上げます。北村大地講師には、論文執筆や研究に関する議論など、細部にわたるまで丁寧にご指導いただきました。特に、ギターの演奏方法に関する知識及びギターリフの魅力を熱く語っていただきました。私自身もギターに興味を持つことができました。ここに感謝申し上げます。

本論の副査である柿元健准教授には、論文の構成や記述に関して大変有益な助言を頂き、大変お世話になりました。ここに感謝申し上げます。

北村研究室の先輩である川口翔也氏、蓮池郁也氏、溝渕悠朔氏、村田佳斗氏には、研究に関するアドバイス等をはじめ、数々のご支援をいただきました。ここに感謝申し上げます。

最後になりますが、現在に至るまで私の学生生活を支えて下さった両親には大変お世話になりました。これまで本当にありがとうございました。最大限の感謝を申し上げます。

参考文献

- [1] A. Bennett, ““Heritage rock”: Rock music, representation and heritage discourse,” *Poetics*, vol. 37, no. 5–6, pp. 474–489, 2009.
- [2] P. Friedlander, *Rock And Roll: A Social History*, New York, Routledge, 2006.
- [3] D. P. Kingma and M. Welling, “Auto-encoding variational bayes,” *Proc. International Conference on Learning Representations*, 2014.
- [4] R. Yu, “A tutorial on VAEs: from bayes’ rule to lossless compression,” arXiv: 2006.10273, 2020.
- [5] A. Roberts, J. Engel, C. Raffel, C. Hawthorne, and D. Eck, “A hierarchical latent vector model for learning long-term structure in music,” *Proc. International Conference on Machine Learning*, vol. 80, pp. 4364–4373, 2018.
- [6] D. P. Kingma, D. J. Rezende, S. Mohamed, and M. Welling “Semi-supervised learning with deep generative models,” *Proc. Neural Information Processing Systems*, pp. 3581–3589, 2014.
- [7] P. Sarmiento, A. Kumar, C. Carr, Z. Zukowski, M. Barthelet and Y. Yang “DadaGP: a dataset of tokenized guitarpro songs for sequence models.” *Proc. International Society for Music Information Retrieval*, pp. 610–617, 2021.
- [8] D. D. Lee and H. S. Seung, “Learning the parts of objects by non-negative matrix factorization,” *Nature*, vol. 401, no. 6755, pp. 788–791, 1999.
- [9] D. D. Lee and H. S. Seung, “Algorithms for non-negative matrix factorization,” *Proc. Neural Information Processing Systems*, pp. 556–562, 2000.
- [10] Guitar pro, <https://www.guitar-pro.com/>
- [11] “Greatest rock guitar riffs 65 essential riffs from legendary songs,” California, Alfred music, 2015.
- [12] Free guitar riffs, <https://www.tabs4acoustic.com/en/free-riffs/guitar-riffs.html>
- [13] 30 guitar riffs that are fun to play & sound awesome, <https://nationalguitaracademy.com/guitar-riffs/>

発表文献一覧

国内学会

1. 唐渡昂希, 北村大地, “ギタータブ譜からのギターリフ抽出アルゴリズムの検討,” 第 25 回 日本音響学会関西支部 若手研究者交流研究発表会, p. 11, 京都, 2022 年 11 月

受賞

1. 第 25 回 日本音響学会関西支部 若手研究者交流研究発表会 奨励賞