

Title	類似フレーズに着目したピアノ演奏の自動採譜
Author(s)	寺井, 浩司
Citation	
Issue Date	1998-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/1166
Rights	
Description	Supervisor:堀口 進, 情報科学研究科, 修士

修 士 論 文

類似フレーズに着目したピアノ演奏の自動採譜

指導教官 堀口 進 教授

北陸先端科学技術大学院大学
情報科学研究科情報システム学専攻

寺井 浩司

1998 年 2 月 13 日

要 旨

本論文では、類似フレーズに着目したピアノ演奏の自動採譜法を提案する。この手法では、曲中に登場する同一あるいは類似のメロディを考慮して採譜を行う。この手法を用いて、人間によるピアノ演奏に含まれるテンポ変動に対処した自動採譜システムの構築を行う。これにより、曲中に登場する同一あるいは類似のメロディ部分の採譜の精度向上を目指し、自動採譜システム全体の精度向上を図る。

目次

1 序論	1
1.1 音楽情報処理	1
1.2 本研究の目的	3
1.3 研究方法	3
1.4 本論文の構成	3
2 自動採譜	4
2.1 はじめに	4
2.2 計算機による自動採譜	4
2.3 電子楽器を利用した自動採譜システム	5
2.4 ビートの追跡	6
3 類似フレーズに着目した自動採譜システム	8
3.1 はじめに	8
3.2 メロディを考慮した自動採譜	8
3.3 フレーズ	9
3.3.1 類似フレーズ	9
3.3.2 同一フレーズ	11
3.4 類似フレーズに着目したピアノ演奏の自動採譜システム	12
3.4.1 自動採譜システムの構成	12
3.4.2 対象とする演奏	14
3.4.3 対象とする楽譜	15
3.4.4 作成される楽譜	15

3.4.5	システム内のデータの流れ	15
3.4.6	ピアノ演奏の収録	17
3.4.7	小節区切れ位置の推定	20
3.4.8	楽譜データの作成	33
3.4.9	フレーズの検査	33
3.4.10	MIDI データの再生成	42
3.4.11	ソフトウェアによる楽譜作成	42
4	自動採譜システムの性能評価	45
4.1	題材とした演奏	45
4.2	小節区切れ位置の正しさ	45
4.3	同一フレーズの抽出率	47
4.4	類似フレーズの抽出率	49
4.5	自動採譜システムの改良	50
4.5.1	小節内のピッチの上下動に関する類似性の利用	50
4.5.2	新しい類似フレーズの具体例	50
4.5.3	新しい類似フレーズ導入の効果	52
5	結論	53
5.1	まとめ	53
5.2	今後の課題	54
	謝辞	55
	参考文献	55
	研究業績	58

第 1 章

序論

1.1 音楽情報処理

近年、計算機や電子楽器の発達および普及、MIDI (Musical Instrument Digital Interface) 規格の登場などにより、計算機で音楽を容易に扱うことが可能になった。これにより、計算機を利用したさまざまな音楽活動が積極的に行われている [1]。その活動は演奏の記録や再生だけでなく、音の加工や生成、自動演奏や自動伴奏、さらに計算機による作曲や編曲など多岐に渡っている。以下に、音楽情報処理について、いくつかの具体例を紹介する。

音の合成

計算機による音の生成に関する研究は、古くから数多く行われている [2]。

初期のものとして、まず、加算方式による音響合成法が挙げられる。これは、実際の音を分析し、その特徴パラメータに基づいて音を合成する分析／再合成方式をとり、代表的なものとして、サイン合成が挙げられる。

つぎに、周波数変調を用いた合成方式である FM 方式 [3] が登場した。これは、市販の音源として広く普及した。

その後、本物の楽器の音をデジタル・サンプリングによりメモリ上に記憶しておき、電子楽器の演奏に対応する音をメモリからロードして再生する PCM 方式が開発され、その音の本物らしさから、FM 音源に代わり市販の音源として普及していった。PCM 方式の問題点は、楽器の音色が無限に存在する一方で、メモリ量の問題から、その一部しかデータとして持つことができず、その再現能力に限界があることである。

また最近の動向として、実際の楽器の音響振動を物理モデルとして計算機でシミュレートする研究が行われている。

自動演奏

計算機による自動演奏として、まず、シーケンサーによる演奏が挙げられる。電子楽器間の音楽データの転送プロトコルを統一した MIDI 規格の登場により、計算機で音楽演奏を容易に扱うことができるようになり、シーケンサーのソフトウェアも多数商品化されている。

自動演奏に関する研究として、楽譜データに表情付けをして人間のような情緒ある演奏を行わせようとする研究が田口 [4] や、福岡ら [5] などによって行われている。また、玉城らは、実際の演奏から抽出した演奏ルールに従って、コンピュータが演奏を再合成する研究を行っている [6] [7]。

楽譜生成

楽譜生成システムは現在実用化されており、すでにかなり質の高い楽譜を生成することができる。読みやすい楽譜を生成するためには、経験や感性に基づく高度な技術が必要であり、現段階では最終的に人間による微調整が必要といわれているが、専門家でない者から見れば、すでに十分質の高い楽譜が得られるように思われる。楽譜生成に関する研究としては、岡が、計算機で楽譜を生成するシステムに関する研究を行っている [8]。

楽譜認識

印刷された楽譜を計算機で認識する研究が行われている。楽譜の認識は一見簡単なように思われるが、実際の楽譜にはかなりの自由度があり、その認識を困難にしている。音楽的な知識を用いて認識精度を向上させる研究が大照 [9] などにより行われている。

自動採譜

音楽演奏を楽譜情報へと変換する技術を自動採譜という。各音のピッチや発音時刻、拍などの情報を演奏から抽出し、楽譜を作成する。計算機への音楽演奏の入力方法としては、演奏の音響信号を入力する方法と、電子楽器を用いて収録した演奏情報を入力する方法がある。音響信号が入力された場合、周波数解析によるピッチ抽出を行う必要がある。

音響信号からの演奏情報の抽出や拍の認識の技術は、まだ改善の余地があり、その意味でも、これらの技術を必要とする自動採譜は、発展途上の段階にあるといえる。

これらの他にも、計算機による作曲および編曲、新世代楽器や指揮システムの研究など、計算機の特徴を活かしたさまざまな研究が行われている。

1.2 本研究の目的

本研究の目的は、類似フレーズに着目したピアノ演奏の自動採譜を行うシステムの構築にある。「多くの曲では、曲中に同一あるいは類似のメロディが繰り返し登場する」という音楽的な性質をビートの追跡に利用する。演奏を小節毎に区切った部分的な演奏の集合として捉え、それらの類似性をチェックしながら自動採譜を行う手法を提案する。これにより、人間による演奏に含まれる微妙なテンポ変動に対処した自動採譜を行うことを目指す。

1.3 研究方法

曲中に同一の部分的なメロディを含む単純な楽曲を用意し、電子ピアノを用いて人間の演奏を MIDI データとして収録する。MIDI データを参照し、小節毎のメロディについて、その類似性をチェックしながら、楽譜データの作成を行う。そして、楽譜作成ソフトを用いて、実際の楽譜に仕上げる。提案した手法で実験を行い、得られた楽譜を評価し、システムの性能について検討する。

1.4 本論文の構成

本論文では、まず、第 2 章で自動採譜システムについて述べる。第 3 章では、類似フレーズに着目した自動採譜法を提案する。第 4 章では、人間によるピアノ演奏を用いて、類似フレーズに着目した自動採譜システムを評価する。そして、最後に第 5 章がまとめとなる。

第 2 章

自動採譜

2.1 はじめに

この章では、自動採譜について説明し、人間による演奏を自動採譜システムで採譜する際の問題点について検討する。

2.2 計算機による自動採譜

楽譜は音楽の記録および伝達における伝統的かつ代表的な手段であるが、採譜作業は専門技術を要する困難な作業であり、その自動化が求められてきた。近年の計算機や電子楽器の発達および普及、また MIDI 規格の登場により、計算機を利用した様々な音楽活動が身近に行われるようになっており、採譜作業についても、計算機を用いた自動採譜システムが実現されている。

計算機への音楽の入力は、大きく分けて 2 通りの方法がある。その 1 つは、演奏をその音響信号として計算機に入力する方法である。もう 1 つは、電子楽器などで演奏を行い、その演奏信号を計算機に入力する方法である。人間は、音楽を聴いて主だった楽器の音色を聞き分け、メロディの流れを追ったり、また音の重なり的美しさを味わったりすることができる。しかし、計算機は人間のように演奏を音楽的に聴くことができないので、演奏をその音響信号として与えられた場合、その音響信号から演奏情報を抽出しなければならない。演奏の音響信号化（デジタル化）自体は難しいことではないが、デジタル化された音楽音響信号から、音楽を表現する基本的な要素である「音」を抽出することは難し

い問題であり、さまざまな研究が行われている [12][13]。一方、電子楽器などで演奏を行い、その演奏信号を計算機に入力する方法では、演奏情報を抽出する必要はなく、計算機は演奏信号を利用してすぐに楽譜作成の作業に入ることができる。

2.3 電子楽器を利用した自動採譜システム

計算機で音楽を扱うための世界的な標準規格として、MIDI が広く普及している。MIDI は、楽器の種類や、発音時刻、ピッチなどのさまざまな演奏情報の記述を定めたものである。MIDI に対応した電子楽器では、MIDI データを用いて自動演奏を行うことができ、また逆に、演奏を直接 MIDI データとして収録することが出来る。

電子楽器で収録した MIDI 形式の演奏データを利用して、計算機により自動採譜を行うことができる。MIDI データを用いた自動採譜システムのソフトウェアとして、Macintosh などで動作する Coda Music Technology の FINALE が挙げられる。FINALE は、音符や休符をはじめとする各パーツを、マウスを用いて 5 線譜上の希望の場所に入力していくことで手軽に楽譜を作ることができる楽譜作成ソフトウェアである。利用者は、必要なパーツを 5 線符上に適当におくだけで良く、あとは FINALE が各パーツのバランスのとれた美しい楽譜に仕上げてくれる。調性や拍節の変更など修正も簡単に行える。

FINALE では、入力方式として、マウスで直接指定していく方法の他に、計算機のキーボードを楽器に見立てたリアルタイム入力、そして MIDI データによる入力がある。MIDI データによる入力では、

- 音符長の量子化レベル
- 拍節構造（拍子記号）
- 調性
- 音符長の量子化レベル

などの各パラメータを指定してやると、計算機が自動的に楽譜を作成してくれる。量子化レベルを大きくすると、大雑把な楽譜が作成され、量子化レベルを小さくすると、複雑な楽譜が作成されてしまう。

演奏を楽譜にするためには、各音のピッチや発音時刻、音の長さなどの情報の他に、その演奏のビートを理解する必要がある。FINALE による MIDI データを用いた自動採譜

では、MIDI データ内のテンポおよびビート情報から演奏のビートを計算している。電子楽器で演奏を収録する際、演奏時の内蔵メトロノームのビート情報を MIDI データに記録するので、FINALE は、電子楽器のメトロノームのビートを演奏のビートとすることになる。しかし、もしメトロノームのビートに合わせて演奏を行ったとしても、人間はメトロノームの指し示すビートに正確に合わせたテンポ一定の演奏を行うことができないので、実際のビートと FINALE が参照するビートの間にずれが生じてしまう。よって、FINALE での自動採譜では、小節の区切り位置を間違えるなど、誤った楽譜を出力してしまうという問題が頻繁に発生する。

2.4 ビートの追跡

音楽演奏のテンポを追跡させようというビート・トラッキングの研究が盛んに行われている [14][15][16]。それまでの演奏のリズム・パターンから次のビートを推測していくのがビート・トラッキングの基本的な考え方である。また、それらのリズム・パターンを階層的に捉えたのも多い。図 2.1 に、ビート・トラッキングの模式図を示す。ほかにも、リアルタイムにビートを追跡するものやリズム・パターンの辞書的知識を利用するものなど、さまざまな報告がなされている。計算機のビート追跡の表現の仕方として、演奏に合わせて機械の足がリズムをとるフット・タッピングという興味深い報告もある [17]。

ビート・トラッキングは、演奏のリズム構造を階層的に捉えることにより行われるが、演奏のメロディをその追跡に利用している例はほとんどない。自動採譜の際には、メロディを考慮してビート・トラッキングを行う手法がより適していると考えられる。

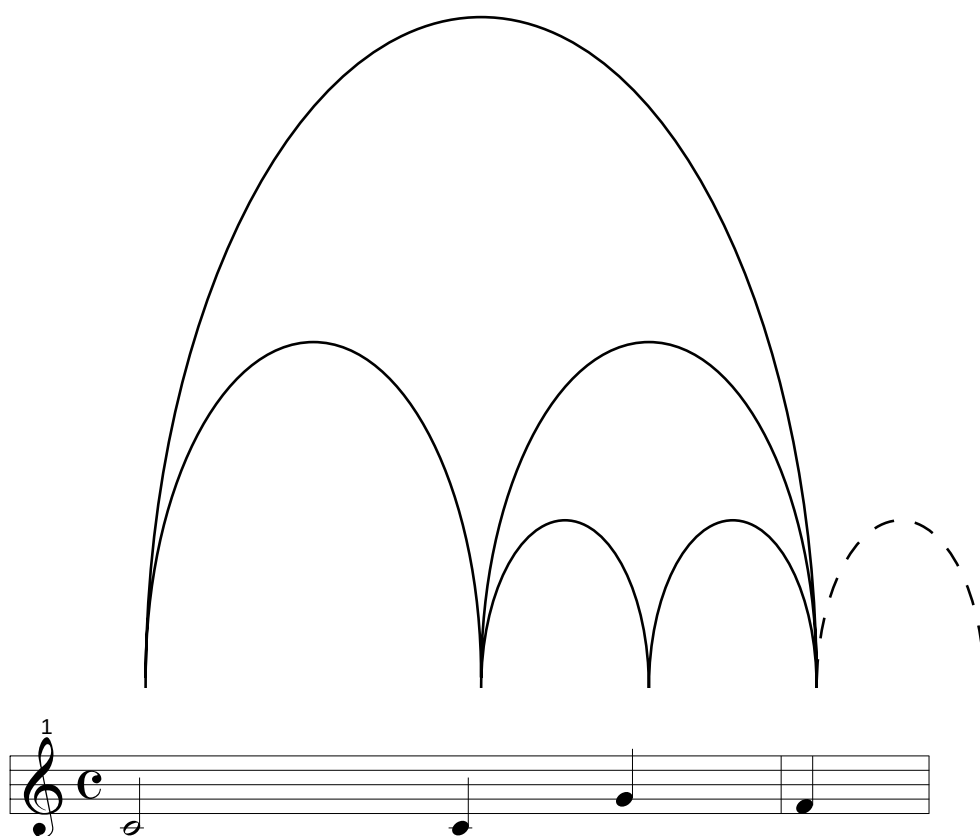


図 2.1: ビート・トラッキング

第 3 章

類似フレーズに着目した自動採譜システム

3.1 はじめに

近年、MIDI 対応の電子楽器を用いて、計算機に演奏情報を入力することが容易になった。電子楽器と計算機の組合せによる自動採譜システムも登場してきている。しかし、自動採譜には、これらの情報の他に、MIDI データには記述されていない、演奏のビートの情報も必要である。MIDI データ内のビート情報では、人間による演奏のビートを正確に表現できておらず、まだ満足できる楽譜を得られないことが多い。ビート・トラッキングは、演奏音の登場するリズムを階層的に捉えそのビートを追跡していく方法だが、演奏音の発音のリズムだけでなく、それらのピッチの情報も利用するメロディとして演奏を捉えたものは少ない。

この章では、演奏中に登場する類似のメロディに注目し、テンポ変動を含む演奏の中でそれらを正しく採譜していく手法として、類似のメロディを考慮した自動採譜を提案する。そして、小節毎の演奏を部分的なメロディの単位とした、類似フレーズに着目した自動採譜法について説明する。

3.2 メロディを考慮した自動採譜

演奏を楽譜にするには、各演奏音のピッチや発音時刻、音の長さといった演奏情報が必要だけでなく、演奏のビート情報を入手しなければならない。そこで、ビート・トラッキングを行ったり、電子楽器に内蔵されたメトロノームの情報を記録した MIDI データ内

のビートおよびテンポ情報で代用したりする。ビート・トラッキングでは、主に、発音リズムの階層的なパターンを演奏から抽出していく方法がとられるが、リズムだけでなく、演奏のメロディも利用する方法も考えられる。MIDI データ内のビートおよびテンポ情報を利用する方法は、便利ではあるが、実際の演奏のビートと異なる危険性が大きい。

ビート・トラッキングは、演奏音の出現のリズムを階層的な構造として捉える手法であり、音楽構造という、音楽のもつ性質の1つを利用しているといえる。演奏と楽譜情報の2つをからめた研究として、カラオケ用の MIDI データについて楽譜構造から演奏を捉え、MIDI データの情報圧縮を行う研究が報告されている [18]。MIDI データと楽譜情報を絡めて演奏を捉えている点で興味深い。

そこで、本論文では、音楽構造を参照してビート追跡に利用することで、テンポ変動に対処した自動採譜を行う手法を提案する。音楽には、「多くの曲は、その曲中に同一あるいは類似のメロディが繰り返し出現する」という性質があることはよく知られている。この、同一あるいは類似のメロディを演奏情報のなかから抽出することで自動採譜の精度の向上を目指す。図 3.1 に、類似のメロディを考慮した自動採譜法の流れを示す。類似のメロディを考慮した自動採譜法では、演奏を分割し、各部分的な演奏を比較し、部分的な演奏間のメロディの類似度を調べ、それらの類似性を採譜に反映させることで、より正しい自動採譜を目指す。

3.3 フレーズ

本論文では、類似のメロディを考慮した自動採譜を行うために、フレーズという演奏単位を用いることにする。楽譜上の1小節分にあたる部分的な楽譜に対する演奏を、フレーズと定義する。類似フレーズに着目した自動採譜法では、小節区切りを推定し、演奏内のフレーズ同士を比較し、フレーズ単位の類似あるいは同一のメロディを演奏から注意深く抽出することで、小節単位の楽譜の類似性を考慮した自動採譜を行う。

3.3.1 類似フレーズ

節単位の部分的な楽譜を曲中から複数取り出し、それぞれの部分的な楽譜内を構成している

- 各音符および休符の種類（長さ）とその順序

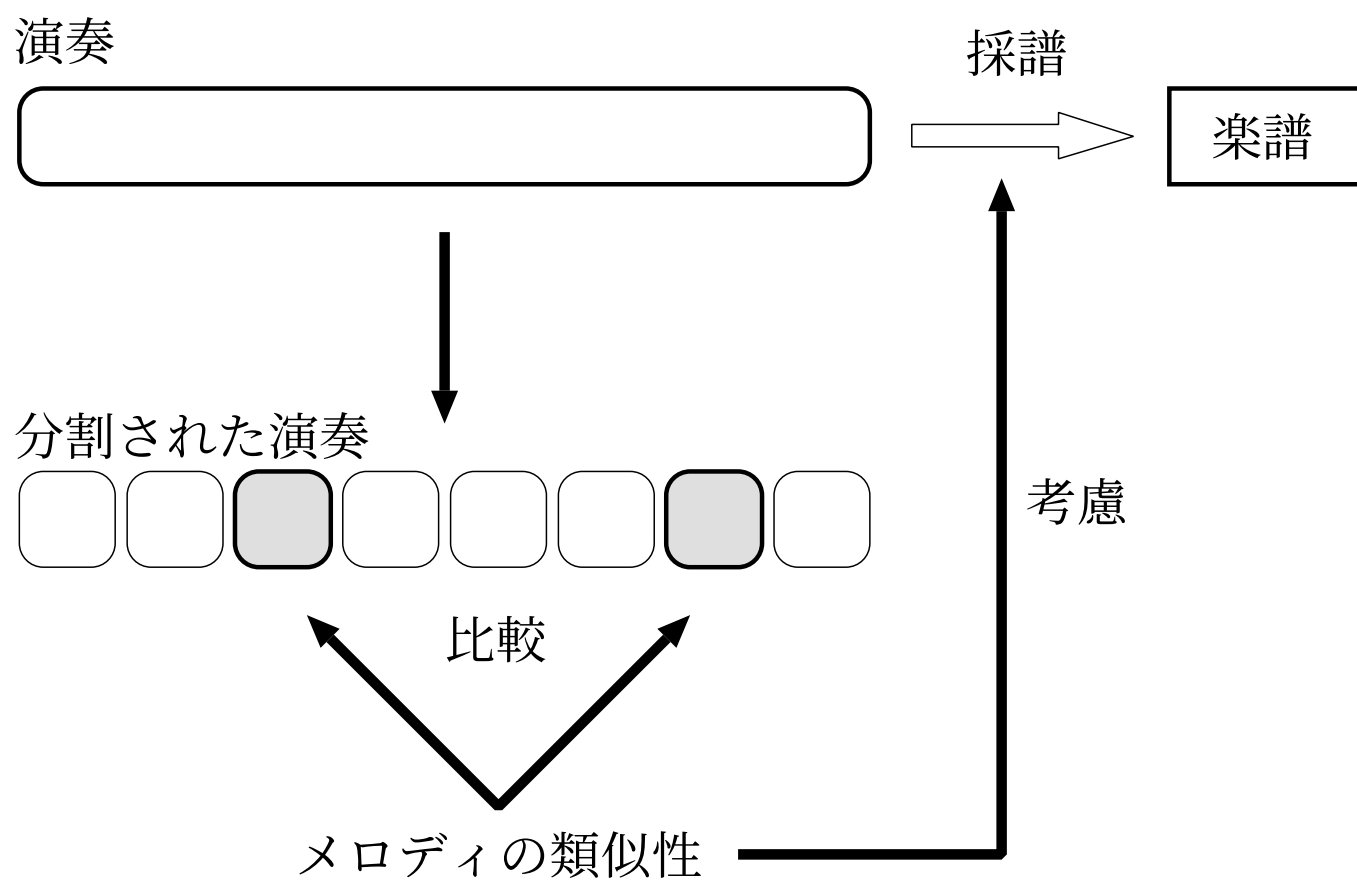


図 3.1: 類似のメロディを考慮した自動採譜法

が等しい時、それらの各部分を演奏したフレーズ同士は類似フレーズの関係にある、と定義する。つまり、小節内の演奏音の発音リズムの等しい小節が複数あり、それらに対する演奏を指して、類似フレーズと呼ぶ。ここで、楽譜では演奏音の発音リズムが等しい小節でも、人間は、楽譜に厳密に正確な演奏を行えないので、発音リズムにある程度の誤差を認めなければならない。類似フレーズの例として、図 3.2 に、「アルプス 1 万尺」の第 1 小節から第 4 小節までの楽譜を示す。この楽譜を演奏した時、第 1 小節、第 2 小節、第 3 小

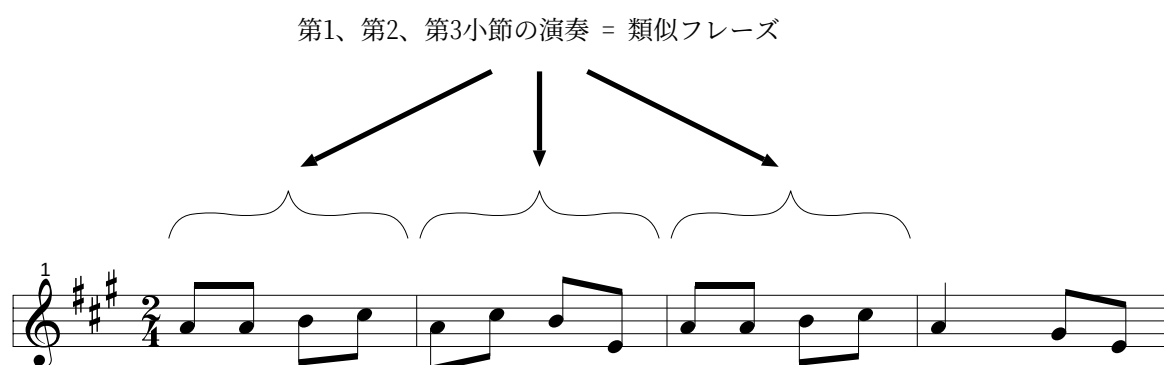


図 3.2: 類似フレーズ

節の各部分的な演奏に対して、これらのフレーズは類似フレーズの関係にある、という。

3.3.2 同一フレーズ

小節単位の演奏を複数取り出してきて、それぞれの部分的な楽譜内を構成している

- 各音符および休符の種類（長さ）とその順序
- 各音符のピッチとその順序

が等しい時、それら部分的な楽譜の演奏であるフレーズ同士は同一フレーズの関係にある、と定義する。同一フレーズは類似フレーズに含まれ、さらに各演奏音のピッチの並びが等しいものを指す。つまり、リズムが類似なものを類似フレーズと呼ぶのに対し、同一フレーズとは、メロディが類似なフレーズの集まりを指している。同一フレーズの例として、図 3.3 に、「アルプス 1 万尺」の楽譜の第 1 小節から第 4 小節までを示す。ここでは、

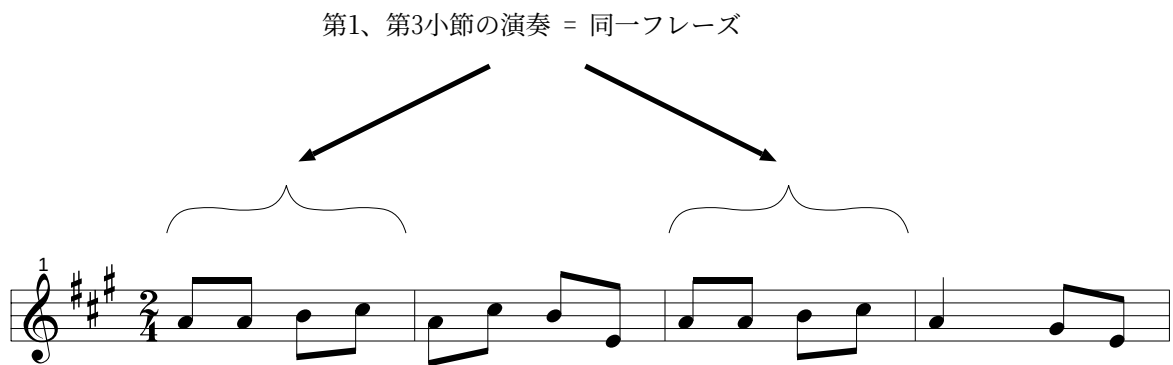


図 3.3: 同一フレーズ

第 1 小節、第 3 小節部分の演奏同士が同一フレーズの関係にある。

3.4 類似フレーズに着目したピアノ演奏の自動採譜システム

本論文では、類似のメロディを考慮した自動採譜として、類似フレーズに着目したピアノ演奏の自動採譜システムを構築する。

3.4.1 自動採譜システムの構成

計算機および MIDI 収録に対応した電子ピアノ、そして楽譜作成ソフトウェアを用いて、類似フレーズに着目したピアノ演奏の自動採譜システムを構築する。図 3.4に、類似フレーズに着目したピアノ演奏の自動採譜システムの構成を示す。演奏者は、採譜を行いたい演奏を電子ピアノで演奏し、MIDI データとして収録する。MIDI データから小節の区切れ位置を推定し、楽譜データを作成する。そして、各フレーズを検査し、類似フレーズの抽出および類似フレーズ部分の楽譜の修正を行う。楽譜データから MIDI データを再生成し、最後に、楽譜作成ソフトウェアの MIDI 入力による楽譜作成機能を利用して楽譜を作成する。

電子ピアノ

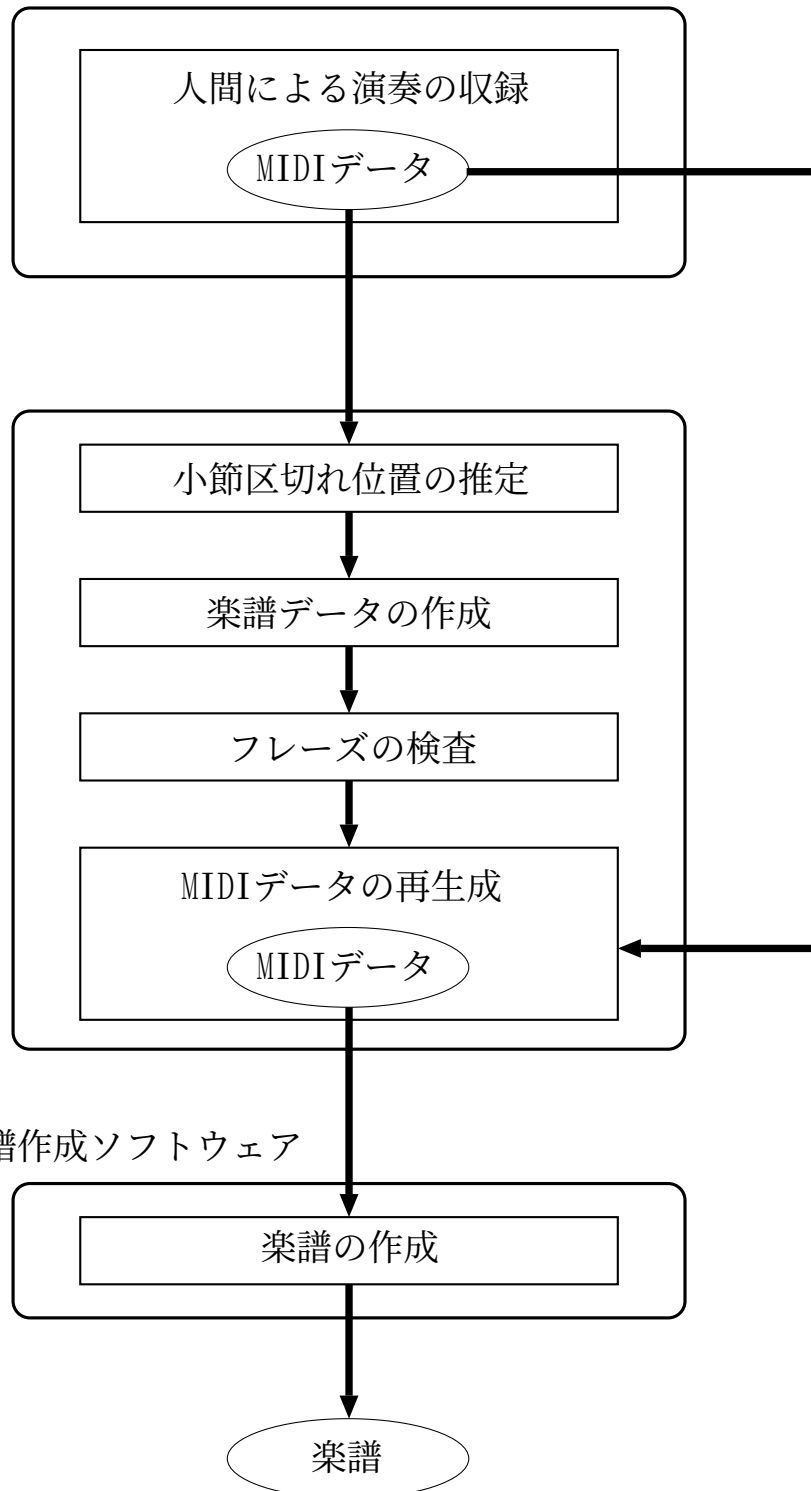


図 3.4: 類似フレーズに着目したピアノ演奏の自動採譜システム

ピアノ演奏の収録

MIDI 対応の電子ピアノを用いて、採譜を行いたい演奏を、MIDI データとして収録する。

小節区切れ位置の推定

ピアノ演奏を収録した MIDI データから、第 1 小節部分のフレーズを仮定し、小節毎の演奏時間の変動を考慮しながら、演奏全体に渡ってフレーズの切り出しを行う。これにより、演奏を楽譜にする際の、小節の区切れ位置の推定を行う。

楽譜データの作成

推定された各小節の区切れ位置をもとに、各小節について、小節内の演奏音のピッチとその発音位置とからなる小節毎の楽譜を作成する。

フレーズの検査

小節毎の楽譜データを比較し、各小節内の構成音のピッチおよび発音位置のパターン（リズム）を比較し、類似フレーズおよび同一フレーズを抽出する。類似フレーズおよび同一フレーズの関係にある各グループの小節内の構成音の発音位置のパターンを統一する。

MIDI データの再生成

フレーズ検査を行った楽譜データと、演奏を収録したものの MIDI データを用いて、楽譜データと同一の MIDI データを作成する。

楽譜の作成

楽譜作成ソフト FINALE の MIDI 入力機能を用いて、再生成された MIDI データから、楽譜を自動作成する。

3.4.2 対象とする演奏

本システムでは、演奏者が採譜を目的として、ある程度テンポ一定の演奏を心がけることを想定している。テンポ一定を心がけた演奏として、

- 演奏者が電子ピアノのメトロノームに合わせた演奏
- メトロノームなしで、演奏者がテンポ一定を心がけた演奏

の2通りの演奏が考えられる。このため、本システムの評価実験では、これら2通りの演奏を用いて行うことにする。また、これらとは別に、プロの演奏者による、一般的な演奏についても、評価実験を行うことにする。

3.4.3 対象とする楽譜

本システムでは、童謡などの単純な楽曲で、

- 楽譜中に和音を含まない、単音演奏（主旋律のみの演奏）の楽譜
- 第1小節の1拍目に発音のある（演奏が始まる）楽譜
- 原則として、各小節の1拍目に発音のある楽譜

の性質を持つ楽譜の演奏を対象とする。

また、これら単純な楽曲にしばしば登場する

- 直前の小節の最後の音とタイで結ばれていて、実際には全く発音のない小節

についても扱えるようにする。

3.4.4 作成される楽譜

本システムでは、各演奏音を、次の演奏音が発音されるまで演奏されていると解釈する。そのため、作成される楽譜には、休符は登場しない。必要な休符は後から手作業で修正して挿入することにする。

3.4.5 システム内のデータの流れ

図3.5に、本システム内のデータの流れを示す。演奏を収録したMIDIデータから楽譜データを作成する。楽譜データの段階でフレーズの検査を行い、その後、楽譜データを表現するMIDIデータを再生成し、それを楽譜作成ソフトウェアに入力することにより、実際の楽譜を得る。

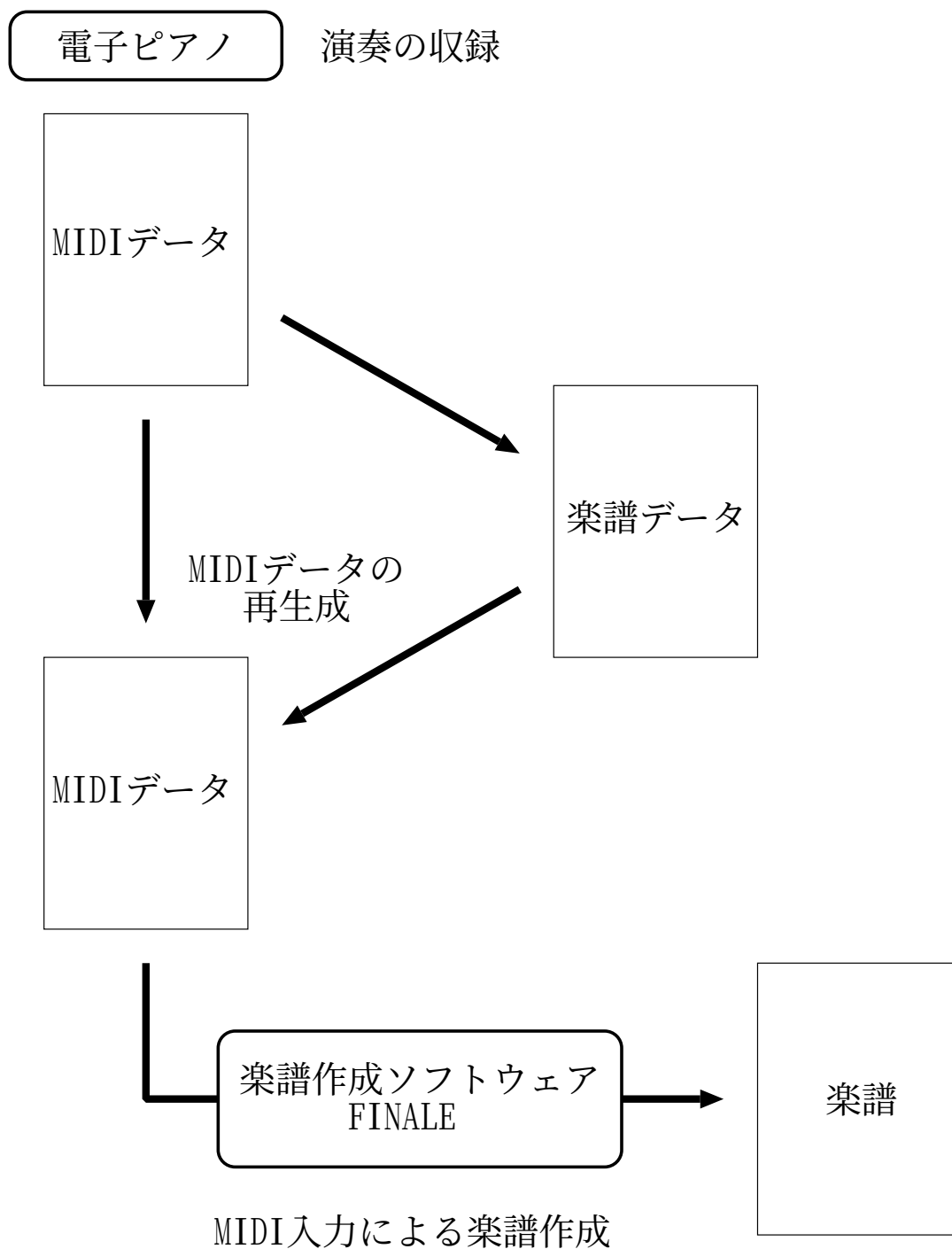


図 3.5: データの流れ

MIDI データ

図 3.6 に、電子ピアノによる演奏を収録した MIDI データをテキスト表示したものの一部を示す。MIDI データには、演奏テンポや拍子記号（ビート）、各演奏音の発音時刻、ピッチ、ヴェロシティや演奏時間、そしてペダル操作などの情報が記録されている。

楽譜データ

MIDI データをもとに作成するものが楽譜データである。楽譜データとは、小節毎に、各演奏音のピッチおよび音符長が記録されたものである。本システムでは休符のない単音演奏の楽譜を作成するので、実際には、各演奏音の音符の長さではなく、小節内での発音の位置が記録されている。図 3.7 に、楽譜データの例を示す。

楽譜

楽譜作成ソフトウェア FINALE の MIDI 入力を利用して、楽譜データから実際の楽譜を作成する。

3.4.6 ピアノ演奏の収録

電子ピアノ YAMAHA Clavinova CVP-83 を用いて、演奏を MIDI データとして収録する。電子ピアノに内蔵されたメトロノームを使用することにより、演奏時のメトロノーム情報（テンポおよびビート）が MIDI データ内に記録される。収録した MIDI データは、スタンダード MIDI ファイルフォーマットで扱う。

YAMAHA Clavinova CVP-83 で演奏を収録すると、総トラック数 4 の MIDI データを作成する。このうち、トラック 1 にはテンポが、トラック 2 には各演奏音の発音時刻、ピッチ、ヴェロシティおよび演奏時間などが記録される。

MIDI データは、時刻とその時刻での演奏動作の対の集合で表されている。各行の先頭にタイムベースをもとにした時刻が記入され、そのあとにそれぞれの動作が示されている。タイムベースは、MIDI データの冒頭に、

`@TIMEBASE 192`

のように記述され、拍の基準となる音符の演奏時間を表す。ビートおよびテンポは、それぞれ時刻 0 の位置に、

%MIDIFILE ABSOLUTE FORM

```

;*****
;  <MThd>  Header & General Info.      0  @CH  1
;      All Tracks : 4 (Trk)             0  CONT  0   0 ;Bank Select (MSB) 00h
;      Date-Time   : 97/08/27 01:55:40  0  CONT 32   0 ;Bank Select (LSB) 00h
;*****                                0  PROG   0
;*****                                0  CONT   7 127 ;Main Volume (MSB) 7Fh
@MThd                                  0  CONT 10  72 ;Panpot (MSB) 48h
@FORMAT 1                             0  CONT 11 127 ;Expression (MSB) 7Fh
@TIMEBASE 192                         0  CONT 64   0 ;Hold1 (Dumper) 00h
@END                                  0  CONT 67   0 ;Soft Pedal 00h
                                         0  CONT 80  64 ;General Purpose #5 40h
;*****                                0  CONT 91  64 ;External Effect 40h
;  <MTrk>  Header Info.                 0  CONT 93   0 ;Chorus Depth 00h
;      Track No.      : 1                0  ~ +0
;      Block Length   : 20 (byte)         0  A4  54      32
;*****                                88  A4  47      56
@MTrk 1                               184  B4  55      56
0  @BEAT 4/4                           276  C#5 57      96
0  @TEMPO 120                           380  A4  51      88
@END                                    448  C#5 50     104
                                         552  B4  51      92
;*****                                656  E4  52      52
;  <MTrk>  Header Info.                 768  A4  58     160
;      Track No.      : 2                948  B4  51      68
;      Block Length   : 391 (byte)       1044 C#5 53     124
;*****
@MTrk 2
0  @BEAT 4/4
                                         :    :    :    :
                                         :    :    :    :

```

図 3.6: MIDI データ (テキスト表示)

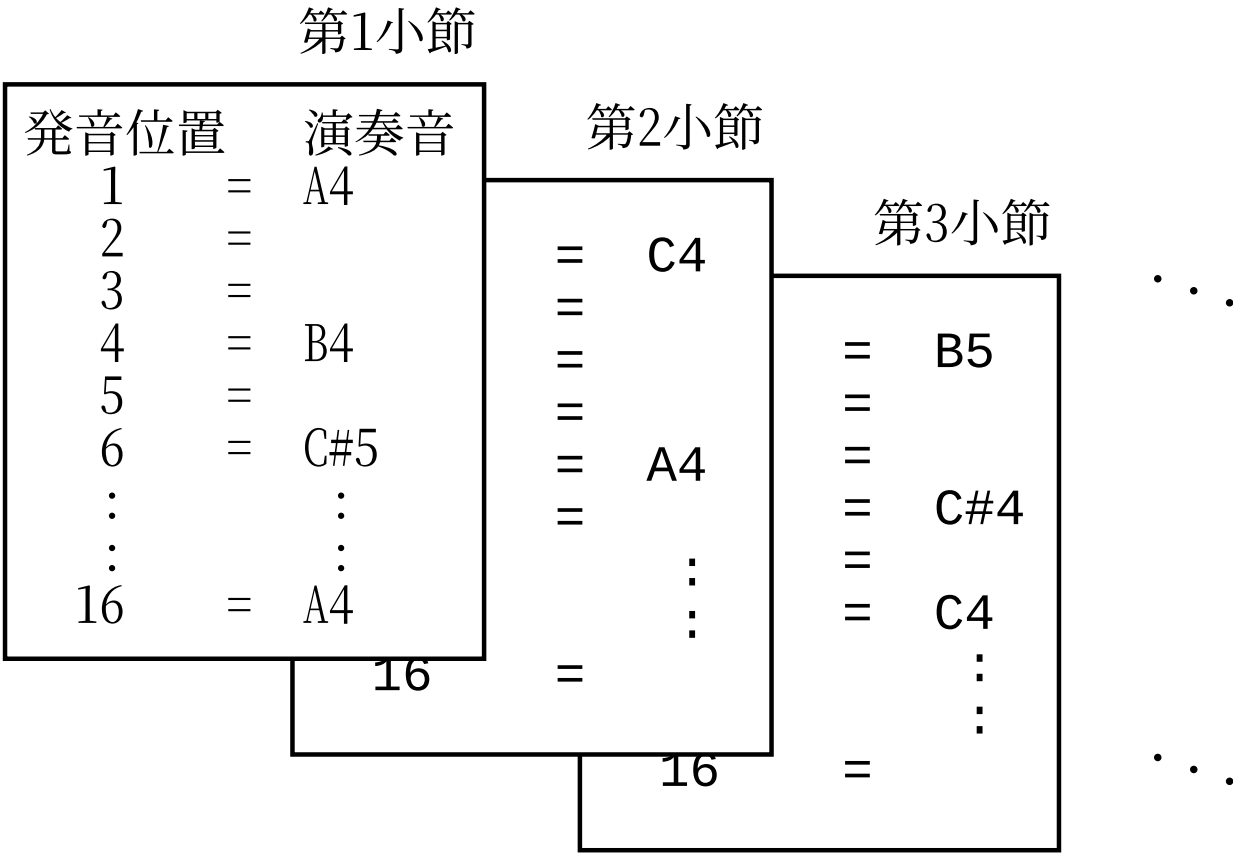


図 3.7: 楽譜データ

0 @BEAT 4/4

0 @TEMPO 120

のように記述される。

各演奏音は、

184 B4 55 56

のように表される。この演奏音は、時刻 184[tick] に発音され、そのピッチは B4 であり、ヴェロシティが 55、演奏時間が 56 で演奏されたことを示している。

i 番目の演奏音の発音時刻 t_i [sec.] は、

$$\begin{aligned} BEAT = b/a & \quad : \text{ビート} \left(\frac{b}{a} \text{拍子} \right) \\ TEMPO & \quad : \text{テンポ} \left(a \text{ 分音符が 1 分間に演奏される回数} \right) \\ TIMEBASE[tick] & \quad : a \text{ 分音符の演奏時間} \end{aligned}$$

であるとする、

$$t_i = \frac{TEMPO}{TIMEBASE \times 60} [sec.] \quad (3.1)$$

で表される。

3.4.7 小節区切れ位置の推定

基本的な作業手順

以下の手順で、MIDI データから、小節区切れ位置の推定を行う。

1. 第 1 小節の構成音数を n 音と仮定する。
2. 1 音目の発音から $(n+1)$ 音目の発音までの時間を計算する。これが第 1 小節の演奏時間 T_1 となる。
3. 第 i 小節の演奏時間 T_i および小節間の演奏時刻変動許容パラメータ d をもとに、第 $(i+1)$ 小節を切り出す。このとき、第 $(i+1)$ 小節の演奏時間 T_{i+1} が

$$\left(1 - \frac{d}{2}\right) T_i \leq T_{i+1} \leq \left(1 + \frac{d}{2}\right) T_i \quad (3.2)$$

を満たすようにする。この条件を満たす時刻に演奏音のない場合は、第 1 小節の構成音数の仮定 n が誤りであるとし、失敗とする。逆にその時刻に複数の演奏音が存在する場合は、そのいずれの切り出し方も成功とする。

4. 演奏の最後まで小節を切り出す。

図 3.8 に、小節区切れ位置推定の様子を示す。この例では、第 4 小節で小節切り出しを失敗し、第 1 小節の構成音数を $n = 2$ とした仮定が誤りとなった。

1 小節分に渡って発音のない場合に対処した方法

直前の小節の最後の発音とタイで結ばれ、1 小節分に渡って発音のない場合に対処した方法を図 3.9 に示す。

小節毎の演奏時間の変動の観察

となり合う小節間の演奏時間の変動許容パラメータ d を決めるために、実験を行った。類似フレーズあるいは同一フレーズを曲中に含む、比較的単純で短い曲として、童謡など以下に示すような曲目を題材とした。

1. アルプス 1 万尺
2. 荒城の月
3. 今日の日はさようなら
4. きよしこの夜

これらの楽曲について、その拍子記号とテンポから計算した、1 小節あたりの演奏時間を調べた。また、各楽曲の小節数と全音数を数え、1 小節あたりの平均の音数を計算した。表 3.1 に、各楽曲について、1 小節あたりの演奏時間と、1 小節あたりの平均の音数を計算した結果を示す。なお、演奏者は、楽譜に記述されたこれらのテンポを目安に演奏を行うものとする。

これらの曲目のうち、例として、図 3.10 に、本実験で使用した「アルプス 1 万尺」の楽譜を示す。

それぞれの曲目について、曲中に類似フレーズおよび同一フレーズがいくつ存在するかを、使用した楽譜で調べた。図 3.11 に、「アルプス 1 万尺」での 3 種類の類似フレーズを、a, b, c の印を付けて示す。また、図 3.12 に、「アルプス 1 万尺」での 2 種類の同一フレーズを A, C の印を付けて示す。

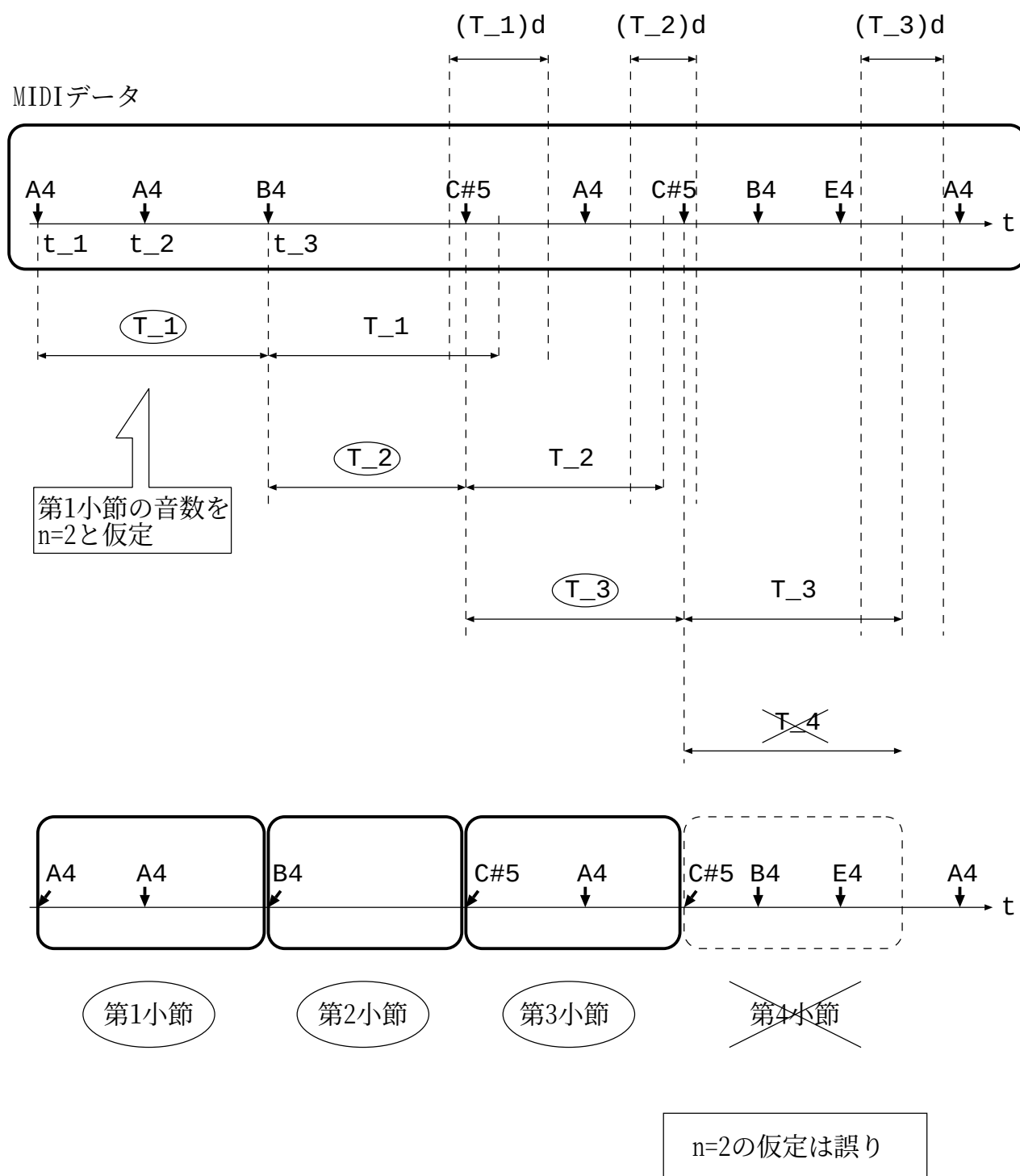


図 3.8: 小節区切れ位置の推定

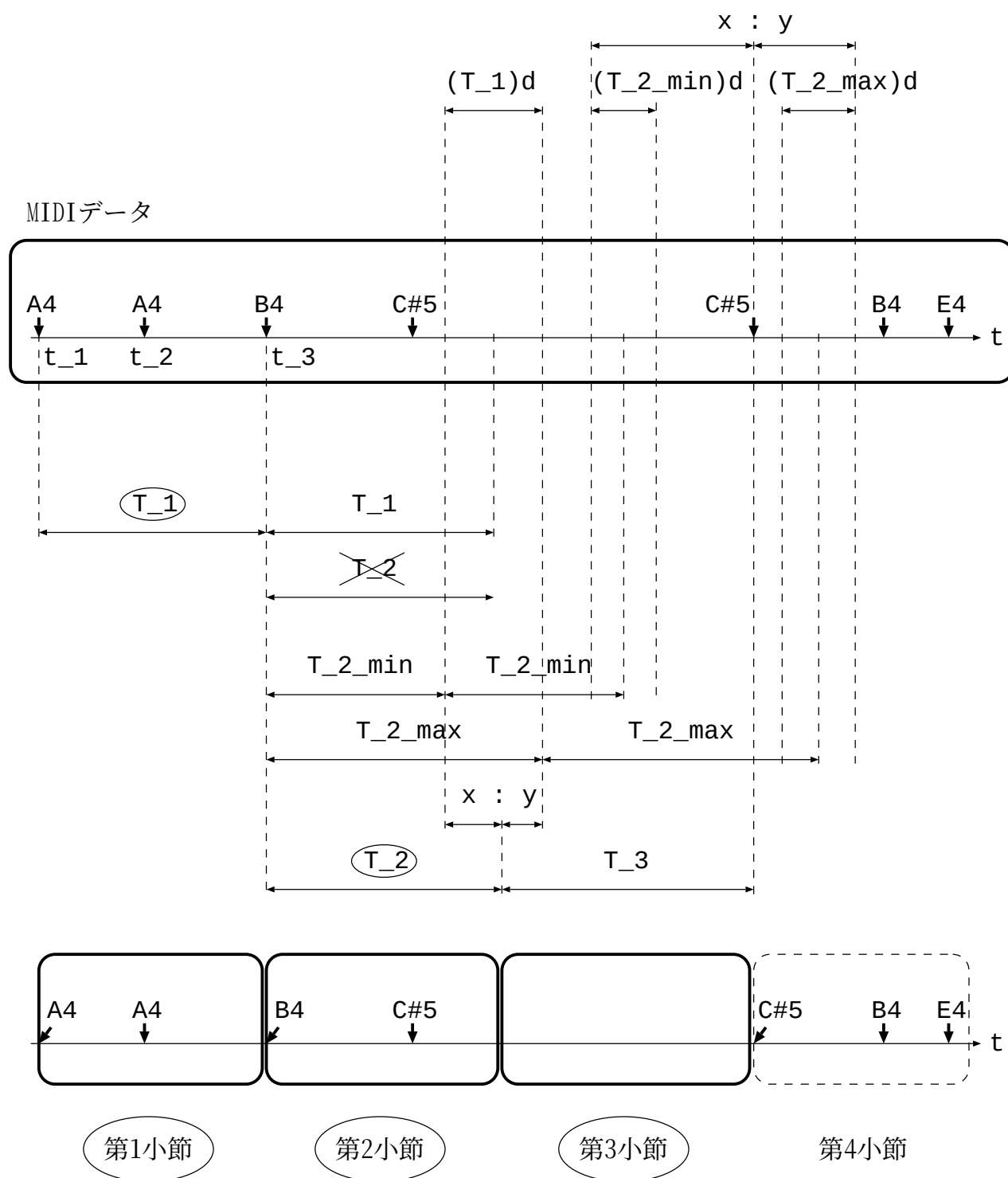


図 3.9: 小節区切れ位置の推定 (1 小節分の無発音への対応)

表 3.1: 題材とした楽曲

曲目	拍子 記号	テンポ	1 小節分の 演奏時間 [sec.]	小節数	音数	1 小節あたりの 音数 (平均)
1. アルプス一万尺	$\frac{2}{4}$	120	1.00	16	57	3.56
2. 荒城の月	$\frac{4}{4}$	70	3.43	16	48	3.00
3. 今日の日はさようなら	$\frac{3}{4}$	80	2.25	16	37	2.31
4. きよしこの夜	$\frac{6}{8}$	80	4.50	12	46	3.83



図 3.10: 「アルプス 1 万尺」の楽譜

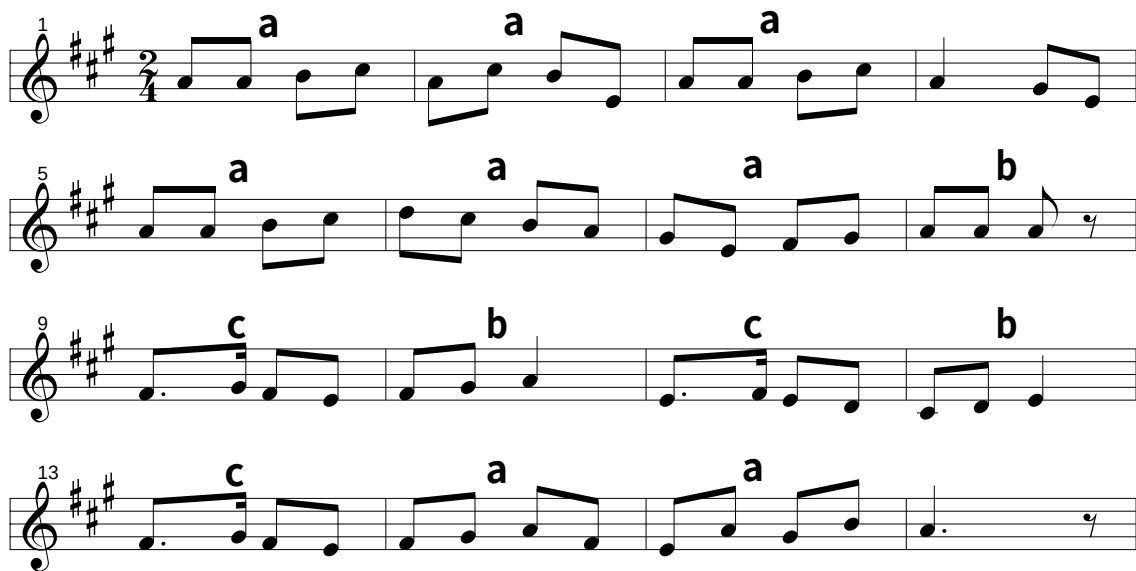


図 3.11: 「アルプス 1 万尺」の類似フレーズ



図 3.12: 「アルプス 1 万尺」の同一フレーズ

ここで、同一フレーズ A は類似フレーズ a に含まれ、同一フレーズ C は類似フレーズ c に含まれている。

これら題材とした曲目について、類似フレーズおよび同一フレーズのそれぞれの個数を楽譜で調べ、表 3.2 に示す。ここで、例えば、「荒城の月」では、

表 3.2: 曲中の類似フレーズ数

曲目	フレーズ数	種類（構成音数）× 個数	
		類似フレーズ	同一フレーズ
1. アルプス一万尺	16	a(4) × 8	A(4) × 3
		b(3) × 3	—
		c(4) × 3	C(4) × 2
2. 荒城の月	16	a(4) × 5	A(4) × 3
		b(3) × 3	B(3) × 3
		c(1) × 4	C1(1) × 2
			C2(1) × 2
		d(4) × 3	D(4) × 2
3. 今日の日はさようなら	16	a(4) × 4	A(4) × 2
		b(3) × 4	—
		c(1) × 2	—
		d(2) × 3	—
4. きよしこの夜	12	a(4) × 4	A(4) × 4
		b(3) × 2	—
		c(5) × 3	C(5) × 2

- フレーズ数（小節数）は、16 である。
- a,b,c,d の 4 パターンの類似フレーズがある。
このうち、例えば類似フレーズ a は、4 音で構成され、曲中に 5 回登場する。
- A,B,C1,C2,D の 4 パターンの同一フレーズがある。
このうち、例えば同一フレーズ C1,C2 は、共に類似フレーズ c に含まれ、また、同

一フレーズ C1,C2 は、1 音で構成され、曲中に各 2 回登場する。

ということを示している。

本研究では、「人間が自動採譜を目的としてなるべくテンポ変動のない演奏を心がけて演奏する」ことを想定している。そのため、以下の条件でそれぞれの演奏を収録した。

- 電子ピアノの内蔵メトロノームに合わせて行った演奏
- メトロノームを使用しない状態で、演奏者がテンポ一定を心がけた演奏

なお、いずれの演奏も、楽譜に記述された演奏テンポを目安にして行われた。それぞれの曲目および演奏の条件において、3 回ずつ演奏を行い、実験に用いた。

図 3.13 に、「アルプス 1 万尺」の、メトロノームに合わせた演奏およびメトロノームを利用しない演奏を各 1 つずつ取り出し、それぞれの小節毎の演奏時間を示す。

いずれの演奏についても、となりあう小節間の演奏時間に変動が見られることが分かる。これは、人間による演奏に必ず含まれるテンポ変動によるものである。メトロノームに合わせた演奏では、小節毎の演奏時間に局所的な変動こそ見られるものの、基本的には、演奏を通して、小節毎の演奏時間は一定であり、演奏のはじめと終わりで小節毎の演奏時間はほとんど変わっていない。一方、メトロノームを利用しない演奏では、演奏を通して、小節毎の演奏時間が変動していることが分かる。となり合う小節間の演奏時間の変動は、大きい所も小さい所も見られる。そして、演奏のはじめと終わりで、小節毎の演奏時間が変化してしまう演奏も見受けられる。

表 3.3 および表 3.4 に、各曲目の、メトロノームに合わせた演奏およびメトロノームを用いない演奏それぞれの、小節毎の演奏時間について示す。

これらの結果から、となり合う小節間の演奏時間の変動の許容パラメータ d を 0.3 とすれば、すべての小節区切れ位置の推定を行えることになる。よって、となり合う小節間の演奏時間の変動の許容パラメータ d を 0.3 として、小節の切り出しを行っていくことにする。

小節区切れ位置推定作業の観察

つぎに、それぞれの演奏について、小節区切れ位置の推定を行った。本実験で題材にした楽曲では、小節内の最大構成音数が 6 なので、それに合わせて、第 1 小節の音数の仮定を 1 から 6 まで行うことにした。また、となり合う小節間の演奏時間の変動許容のパ

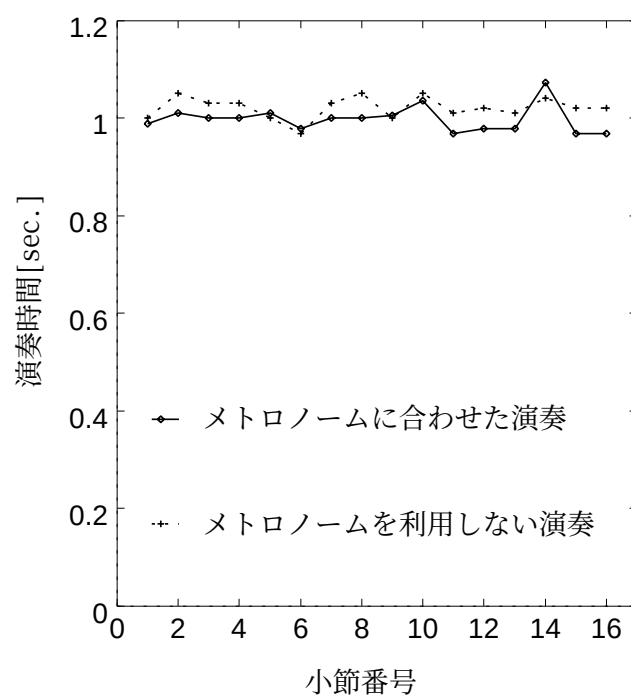


図 3.13: 小節毎の演奏時間—「アルプス 1 万尺」

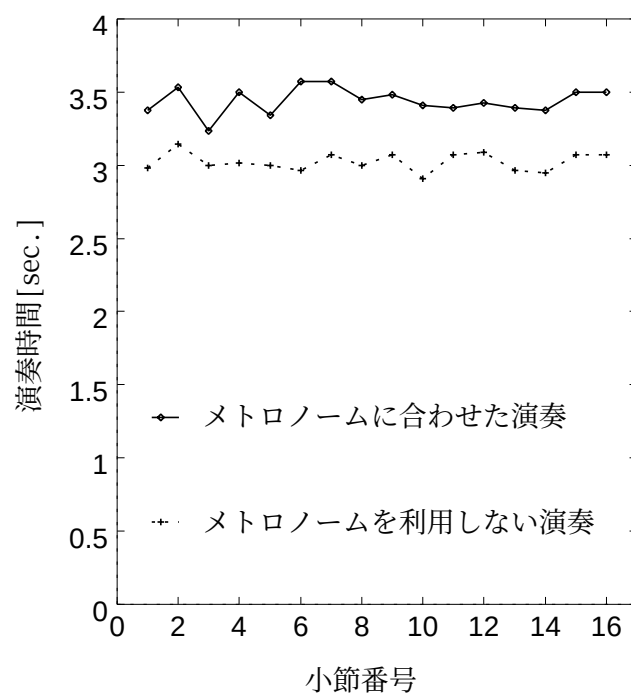


図 3.14: 小節毎の演奏時間—「荒城の月」

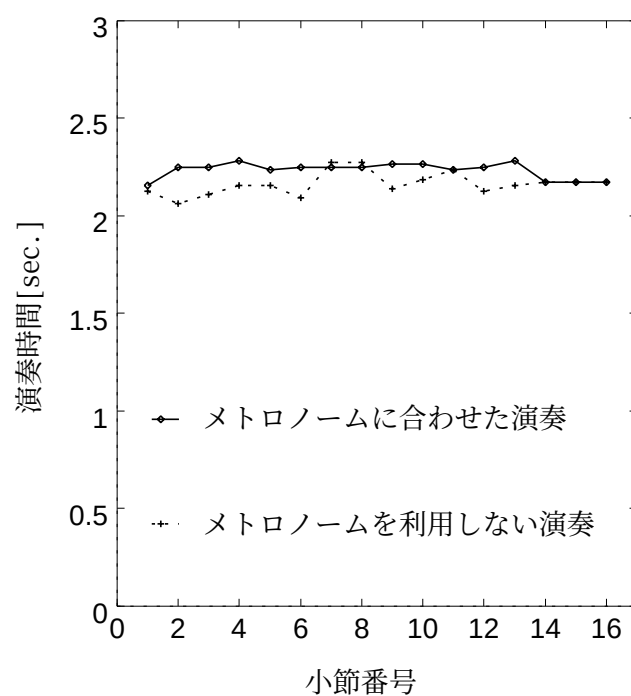


図 3.15: 小節毎の演奏時間—「今日の日はさようなら」

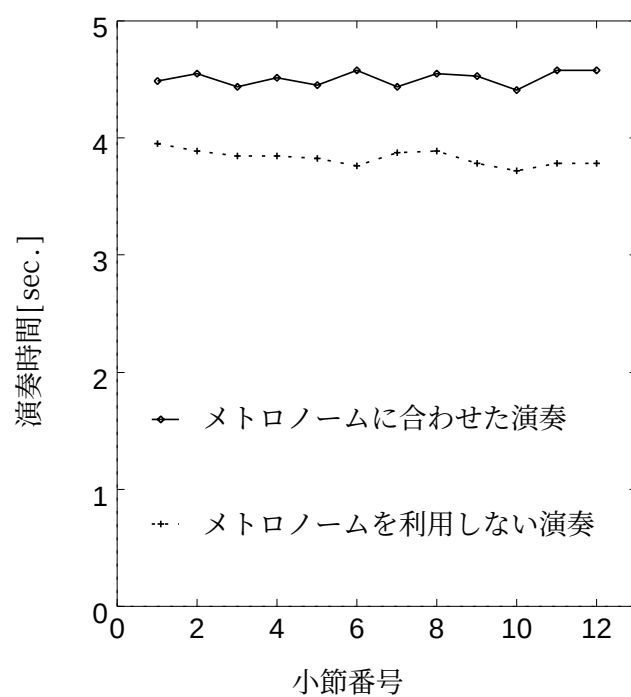


図 3.16: 小節毎の演奏時間—「きよしこの夜」

表 3.3: となり合う小節間の演奏時間の変動割合（その 1）

曲目	テンポ	演奏		第 i 小節の演奏時間 第 $(i-1)$ 小節の演奏時間	
				最大	最小
1. アルプス一万尺	120	メトロノーム 使用	演奏 1	1.10	0.90
			演奏 2	1.12	0.90
			演奏 3	1.07	0.93
			平均	1.10	0.91
		メトロノーム 使用せず	演奏 1	1.06	0.95
			演奏 2	1.07	0.98
			演奏 3	1.05	0.95
			平均	1.06	0.96
		平均		1.08	0.94
2. 荒城の月	70	メトロノーム 使用	演奏 1	1.08	0.91
			演奏 2	1.07	0.93
			演奏 3	1.11	0.90
			平均	1.09	0.91
		メトロノーム 使用せず	演奏 1	1.05	0.95
			演奏 2	1.05	0.93
			演奏 3	1.03	0.94
			平均	1.04	0.94
		平均		1.07	0.93

表 3.4: となり合う小節間の演奏時間の変動割合（その 2）

曲目	テンポ	演奏		第 i 小節の演奏時間 第 $(i-1)$ 小節の演奏時間	
				最大	最小
3. 今日の日はさようなら	80	メトロノーム 使用	演奏 1	1.04	0.95
			演奏 2	1.10	0.92
			演奏 3	1.08	0.93
			平均	1.07	0.93
		メトロノーム 使用せず	演奏 1	1.09	0.94
			演奏 2	1.08	0.94
			演奏 3	1.09	0.93
			平均	1.09	0.94
		平均		1.08	0.94
4. きよしこの夜	40	メトロノーム 使用	演奏 1	1.04	0.97
			演奏 2	1.09	0.89
			演奏 3	1.05	0.96
			平均	1.06	0.94
		メトロノーム 使用せず	演奏 1	1.03	0.97
			演奏 2	1.06	0.90
			演奏 3	1.03	0.96
			平均	1.04	0.94
		平均		1.05	0.94

ラメータは、前述の通り、 $d=0.3$ とした。表 3.5 に、「アルプス 1 万尺」について、小節区切れ位置の推定を行い、得られた小節切り出し方の候補数をまとめた。「アルプス 1 万尺」では、第 1 小節の音数は $n=4$ であるが、 $n=4$ 以外の仮定でも、演奏の小節区切れ位置の推定に成功してしまうことが分かった。また、ある仮定について、複数の切り出し方ができてしまうことがあることが分かった。

同様にして、表 3.6 に、「荒城の月」に対しての、小節の切り出し方の候補数を示す。また、表 3.7 に、「今日の日はさようなら」に対しての、小節の切り出し方の候補数を、表 3.8 に、「きよしこの夜」に対しての、小節の切り出し方の候補数を示す。

いずれの題材についても、1 つの演奏について、複数の小節切り出しの候補が得られることが分かった。また、正しい切り出し方を行った候補が、これらの中に必ず含まれていることが分かった。ここでは小節区切れ位置の推定の手法はこのままにし、その作業の後で、それらの候補を 1 つに絞り込む方法を考えることにする。

第 1 小節の音数を少なく仮定した場合、小節の切り出しを失敗してしまう可能性があることが分かった。これは、実際よりも短い長さで小節を仮定してしまい、それにより、2 小節以上に渡って演奏音の発音のない部分ができてしまい、それに対して、この手法では対応できていないことを示している。しかし、これら切り出しに失敗したものは、いずれも正しくない小節の切り出しについて失敗したものであり、この手法について考え直す必要はないものと思われる。

表 3.17 に、各曲目別に、小節の切り出し方の候補数をまとめたものを示す。

3.4.8 楽譜データの作成

小節毎に分割された部分的な MIDI データを元に、楽譜データを作成する。小節内を等分し、それらの位置で各演奏音の発音時刻を表現する。表 3.9 に、拍子記号と小節内の等分数の関係を示す。

図 3.18 に、1 小節分の MIDI データから楽譜データを作成する例を示す。

3.4.9 フレーズの検査

MIDI データから作成した楽譜データを用いて、フレーズの検査を行う。発音パターンの類似なものを抽出し、類似フレーズあるいは同一フレーズとする。そして、類似フレーズおよび同一フレーズとみなされた各フレーズの発音パターンを統一する。

表 3.5: 小節の切り出し方の候補数—「アルプス 1 万尺」

演奏		第 1 小節の音数						合計
		1	2	3	4	5	6	
メトロノーム 使用	演奏 1	1	1	1	1	1	2	7
	演奏 2	1	2	2	1	2	2	10
	演奏 3	—	2	1	1	2	2	8
	平均	0.67	1.67	1.33	1.00	1.67	2.00	8.33
メトロノーム 使用せず	演奏 1	—	2	1	1	2	2	8
	演奏 2	1	3	1	1	2	2	10
	演奏 3	—	4	2	1	2	2	11
	平均	0.33	3.00	1.33	1.00	2.00	2.00	9.67
平均		0.50	2.33	2.67	1	2.00	2.00	9.00

表 3.6: 小節の切り出し方の候補数—「荒城の月」

演奏		第 1 小節の音数						合計
		1	2	3	4	5	6	
メトロノーム 使用	演奏 1	—	1	1	2	1	1	6
	演奏 2	—	2	2	4	1	1	10
	演奏 3	—	1	1	2	1	1	6
	平均	—	1.33	1.33	2.33	1.00	1.00	7.33
メトロノーム 使用せず	演奏 1	—	1	1	4	2	1	9
	演奏 2	—	2	2	2	1	2	9
	演奏 3	—	3	2	4	2	2	13
	平均	—	2.00	1.67	1.00	1.67	1.67	10.33
平均		0.00	1.67	1.50	3	2.67	1.33	8.83

表 3.7: 小節の切り出し方の候補数—「今日の日はさようなら」

演奏		第 1 小節の音数						合計
		1	2	3	4	5	6	
メトロノーム 使用	演奏 1	—	—	—	4	4	4	12
	演奏 2	—	—	—	2	4	2	8
	演奏 3	—	—	—	2	4	4	10
	平均	0.00	0.00	0.00	3.00	4.00	3.33	10.00
メトロノーム 使用せず	演奏 1	—	—	—	4	4	4	12
	演奏 2	—	—	—	4	8	4	16
	演奏 3	—	—	—	4	4	4	12
	平均	0.00	0.00	0.00	4.00	5.33	4.00	13.33
平均		0.00	0.00	0.00	3.33	4.67	3.67	11.67

表 3.8: 小節の切り出し方の候補数—「きよしこの夜」

演奏		第 1 小節の音数						合計
		1	2	3	4	5	6	
メトロノーム 使用	演奏 1	1	2	3	2	4	4	16
	演奏 2	1	3	4	2	4	4	18
	演奏 3	1	2	4	1	2	2	12
	平均	1.00	2.33	3.67	1.67	3.33	3.33	15.33
メトロノーム 使用せず	演奏 1	1	4	4	1	4	4	18
	演奏 2	1	4	8	1	8	8	30
	演奏 3	1	4	4	2	4	4	19
	平均	1.00	4.00	5.33	1.33	5.33	5.33	22.33
平均		1.00	3.17	4.50	1.50	4.33	4.33	18.83

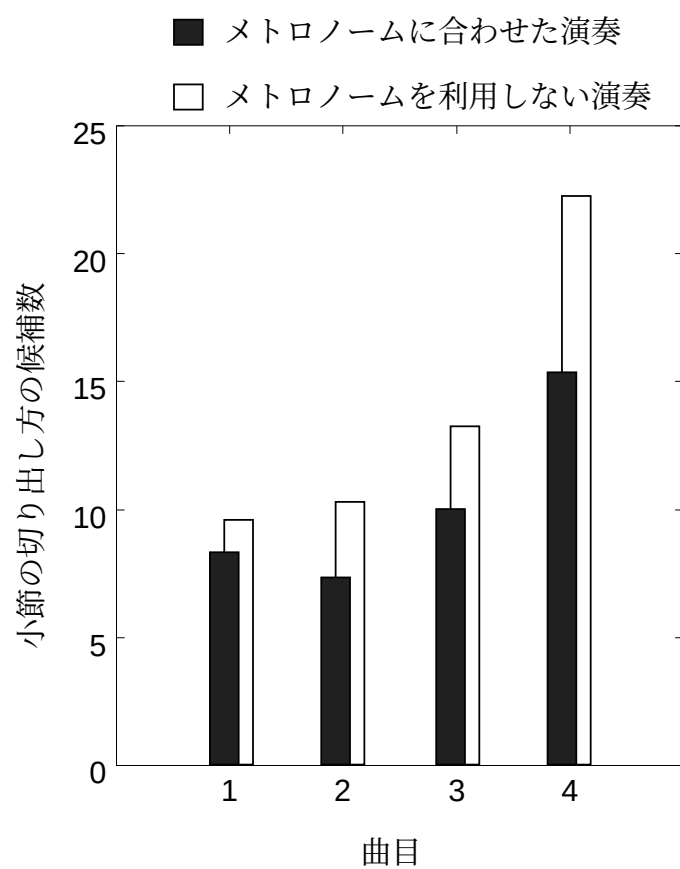
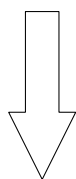
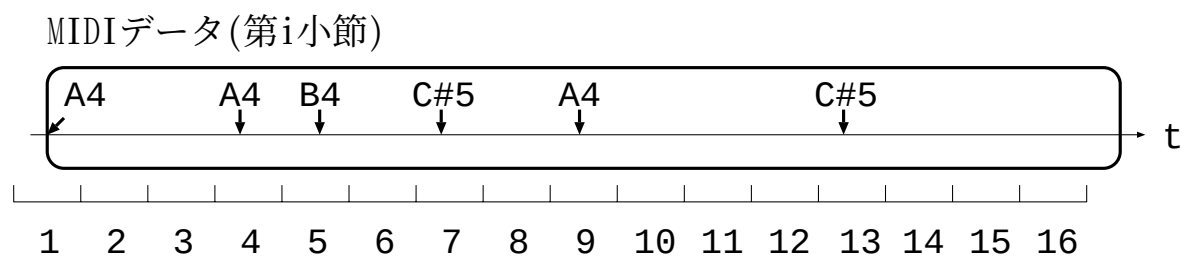


図 3.17: 小節の切り出し方の候補数

表 3.9: 拍子記号と小節内の等分数の関係

分割数	拍子記号
16	$\frac{4}{4}$
	$\frac{2}{4}$
12	$\frac{3}{4}$
	$\frac{6}{8}$



楽譜データ(第i小節)

発音位置	演奏音
1	= A4
2	=
3	=
4	= A4
5	= B4
6	=
7	= C#5
8	=
9	= A4
10	=
11	=
12	=
13	= C#5
14	=
15	=
16	=

図 3.18: 楽譜データの作成

類似フレーズの抽出

人間は楽譜通りに正確なテンポで演奏できないので、例えば類似フレーズ同士の発音パターンにも誤差が生じることが予想される。そこで、正しく小節区切れ位置の推定を行ったものについて、類似フレーズ同士が、実際はどの程度の発音パターンのずれを持っているのかを検査した。まず、「アルプス1万尺」のメトロノームに合わせた演奏を1つとりだし、演奏中に3回登場する同一フレーズ A 内の構成音の発音パターンを調べた。図 3.19に、実際の楽譜での発音パターンと、これら演奏中の各フレーズの発音パターンを示す。同一フレーズ A は、演奏中に3回登場するが、そのいずれも、実際の楽譜での発音

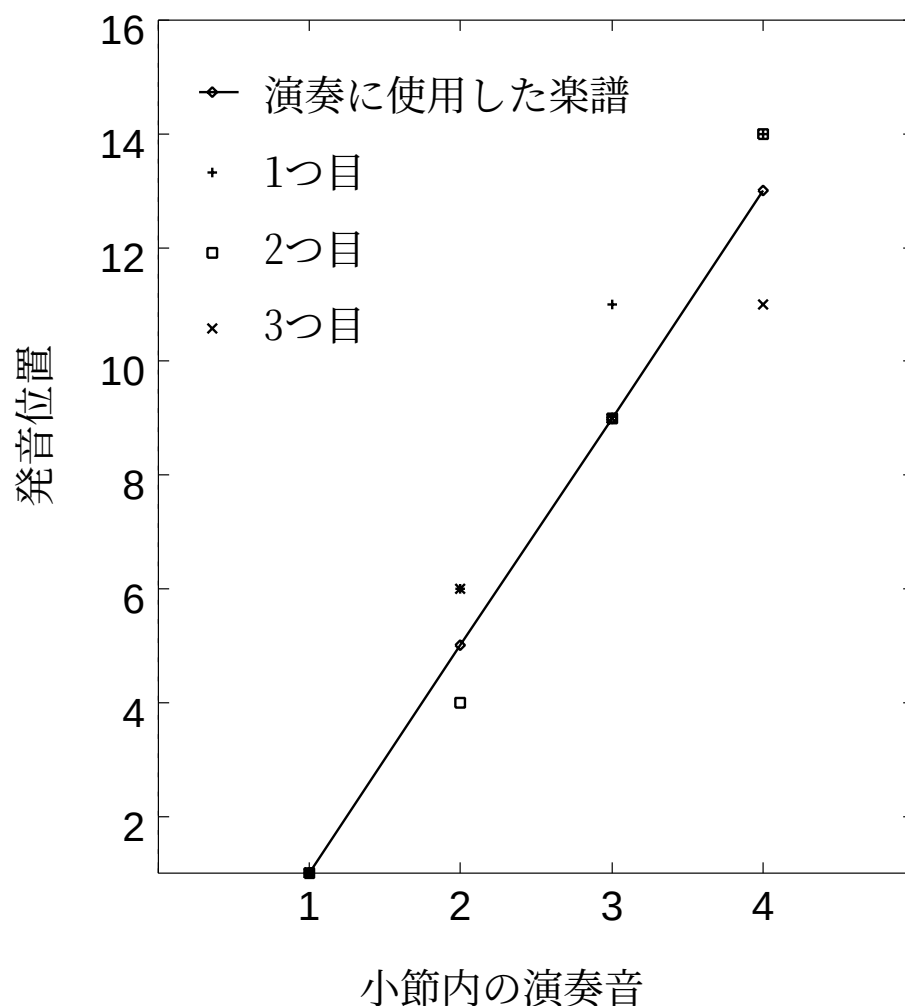


図 3.19: 同一フレーズ内の構成音の発音位置のずれ

パターンとはことなるパターンを持っていることが分かる。これら3つのことなる発音パターンのものを同一フレーズの関係として認めなければならない。このように、類似／同一フレーズを正しく抽出するためには、発音パターンを比較する際、ある程度の誤差を認めてやる必要があることが分かった。

類似／同一フレーズを正しく抽出するために発音パターンの誤差をどの程度認めるべきかを定めるために、各演奏の類似フレーズ内の構成音の発音パターンを調べた。表 3.10、表 3.11 および表 3.12 に、「アルプス 1 万尺」のメトロノームなしの演奏について、各類似／同一フレーズの発音パターンを調べたものを示す。

これら「アルプス 1 万尺」の、各類似／同一フレーズについて、各構成音の発音位置のずれの最大値を調べたものを、表 3.13 に示す。

これから、第 1 音目は、いずれの類似フレーズにおいてもその発音位置に誤差がないことが分かる。これは、本手法では、1 音目を小節の 1 拍目として切り出すため、発音位置が必ず 1 になるためである。2 音目以降に関しては、その発音位置にばらつきがあることが分かる。1 音あたりの誤差の最大は、類似フレーズ b の 2 音目の、4 である。一方、これら小節内の各構成音の発音位置の誤差を足し合わせた合計の誤差を調べたところ、最大値は 6 となった。しかし、類似／同一フレーズ内の構成音の発音位置の合計の誤差を 6 ととってやると、本来類似／同一フレーズの関係にないものまで含んでしまう。

ところで、ここで、同一フレーズの総数を目安にして小節区切れ位置推定の複数の候補を絞ることができることが分かった。例として、表 3.14 に、「アルプス 1 万尺」のメトロノームなしの演奏について、小節の切り出し方の候補と類似フレーズ数の関係を示す。この演奏では、第 1 小節の音数を 4 と仮定した時に、正しい小節区切れ位置の推定を行っているが、この時、同一フレーズの総数が最大値を示している。これは他の演奏や楽曲についても同様なことがいえる。そこで、同一フレーズの総数が最大値を示したものを、正しい小節切り出しを行ったものとして選択することができる。

類似フレーズ部分の楽譜データの統一

類似フレーズの発音パターンの平均を取って、類似フレーズ部分の統一発音パターンとする。この作業で、類似フレーズ部分の楽譜データが統一される。

表 3.10: 類似フレーズ内の構成音の発音パターン—「アルプス 1 万尺」(その 1)

類似フレーズ a			同一フレーズ A		
演奏 1	演奏 2	演奏 3	演奏 1	演奏 2	演奏 3
(1,4,10,14)	(1,4,9,13)	(1,4,11,13)	(1,4,10,14)	(1,4,9,13)	(1,4,11,13)
(1,4,8,12)	(1,6,10,14)	(1,5,9,13)	—	—	—
(1,4,11,14)	(1,5,10,13)	(1,4,11,13)	(1,4,11,14)	(1,5,10,13)	(1,4,11,13)
(1,5,9,13)	(1,5,10,14)	(1,4,10,13)	(1,5,9,13)	(1,5,10,14)	(1,4,10,13)
(1,5,8,12)	(1,4,8,14)	(1,6,10,14)	—	—	—
(1,7,10,14)	(1,7,8,15)	(1,5,10,13)	—	—	—
(1,4,10,15)	(1,4,11,14)	(1,5,11,14)	—	—	—
(1,5,10,14)	(1,7,11,15)	(1,4,8,15)	—	—	—

表 3.11: 類似フレーズ内の構成音の発音パターン—「アルプス 1 万尺」(その 2)

類似フレーズ b		
演奏 1	演奏 2	演奏 3
(1,4,8)	(1,7,10)	(1,4,10)
(1,3,9)	(1,4,10)	(1,4,11)
(1,6,11)	(1,6,10)	(1,4,11)

表 3.12: 類似フレーズ内の構成音の発音パターン—「アルプス 1 万尺」(その 3)

類似フレーズ c			同一フレーズ C		
演奏 1	演奏 2	演奏 3	演奏 1	演奏 2	演奏 3
(1,7,10,13)	(1,6,9,12)	(1,6,9,14)	(1,7,10,13)	(1,6,9,12)	(1,6,9,14)
(1,8,10,13)	(1,9,11,13)	(1,6,11,14)	—	—	—
(1,6,10,14)	(1,7,9,12)	(1,7,10,13)	(1,6,10,14)	(1,7,9,12)	(1,7,10,13)

表 3.13: 類似フレーズ内の構成音の発音位置のずれの最大—「アルプス 1 万尺」

類似フレーズ	1 音目	2 音目	3 音目	4 音目	全体
類似フレーズ a	0	3	3	3	6
同一フレーズ A	0	1	2	1	4
類似フレーズ b	0	4	3	—	5
類似フレーズ c	0	3	2	2	4
同一フレーズ C	0	1	1	2	3

表 3.14: 小節の切り出し方の候補と類似フレーズ数の関係—「アルプス 1 万尺」

小節の切り出し方の候補		小節数	類似フレーズ数	同一フレーズ数
第 1 小節の 音数	3	21	16+2	2
	4	16	8+3+3	3+2
	5	13	6	—
	6	11	5+2	2
		11	4+2	—

3.4.10 MIDI データの再生成

楽譜データそのままの楽譜を FINALE で作成するために、楽譜データと同一の MIDI データを再生成する。

$\frac{4}{4}$ 拍子や $\frac{2}{4}$ 拍子の演奏では、 $BEAT = b/a$ での第 i 小節の発音位置 j の音の発音時刻 $t_{ij}[\text{tick}]$ は、

$$t_{ij} = \left((i - 1) + \frac{j - 1}{16} \right) \times TIMEBASE \times b \quad (3.3)$$

となる。また、 $\frac{3}{4}$ 拍子や $\frac{6}{8}$ 拍子の演奏ならば、

$$t_{ij} = \left((i - 1) + \frac{j - 1}{12} \right) \times TIMEBASE \times b \quad (3.4)$$

となる。

3.4.11 ソフトウェアによる楽譜作成

MIDI 入力により、自動採譜を行う量子化レベルを正しく選び、楽譜データと同一のものを作成する。

図 3.20 に、メトロノームに合わせた「アルプス 1 万尺」の演奏の、本自動採譜システムによる楽譜の出力結果を示す。

ここで、類似フレーズに着目した利点を示すために、図 3.21 に、メトロノームに合わせた「アルプス 1 万尺」の演奏の、フレーズ検査を行わずに作成した楽譜を示す。

2 つの楽譜を比較すると、類似／同一フレーズ部分の部分的な楽譜に違いが現れているのが分かる。

The image displays a musical score for three tracks, labeled Track 1, Track 2, and Track 3. The score is organized into three systems, each beginning with a measure number (1, 5, and 9). The key signature is three sharps (F#, C#, G#) and the time signature is 2/4. Track 1 contains the melody, while Tracks 2 and 3 contain accompaniment parts. The melody in Track 1 consists of eighth and quarter notes, with some measures containing beamed eighth notes. The accompaniment in Tracks 2 and 3 consists of quarter notes and rests.

Track 1

Track 2

Track 3

5

5

5

9

9

9

図 3.20: メトロノームに合わせた演奏の自動採譜結果

Track 1

Track 2

Track 3

5

9

図 3.21: フレーズ検査を行わない自動採譜結果

第 4 章

自動採譜システムの性能評価

4.1 題材とした演奏

これまで題材としてきた 4 つの曲目の計 24 の演奏に加え、演奏家による演奏を用意して、自動採譜システムの評価実験を行った。演奏家による演奏は

- 「きよしこの夜」演奏 1
- 「きよしこの夜」演奏 2
- 「荒城の月」

の 3 つの演奏で、それぞれ主旋律のみを取り出した単音演奏の MIDI データを作成して評価実験に用いた。

4.2 小節区切れ位置の正しさ

表 4.1 に、メトロノームを使用した演奏について、楽曲毎に、小節区切れ位置の正しさを調べて示す。

また、表 4.2 に、メトロノームを使用しない演奏について、楽曲毎に、小節区切れ位置の正しさを調べて示す。

表 4.3 に、演奏家による演奏について、楽曲毎に、小節区切れ位置の正しさを調べて示す。いずれの演奏についても、FINALE と比較して、本手法は、より小節区切れ位置を正

表 4.1: 小節区切れ位置の正しさ（メトロノーム使用）

演奏	FINALE	本手法
「アルプス 1 万尺」	63%	100%
「荒城の月」	56%	100%
「今日の日はさようなら」	58%	96%
「きよしこの夜」	58%	100%

（ 値は、それぞれ 3 つの演奏の平均）

表 4.2: 小節区切れ位置の正しさ（メトロノーム使用せず）

演奏	FINALE	本手法
「アルプス 1 万尺」	23%	100%
「荒城の月」	21%	100%
「今日の日はさようなら」	42%	100%
「きよしこの夜」	17%	97%

（ 値は、それぞれ 3 つの演奏の平均）

表 4.3: 小節区切れ位置の正しさ（演奏家による演奏）

演奏	FINALE	本手法
「きよしこの夜」演奏 1	39%	100%
「きよしこの夜」演奏 2	22%	86%
「荒城の月」	23%	96%

しく推測していることが分かる。しかし、メトロノームを使用した「今日の日はさようなら」の演奏の1つ、そしてメトロノームを使用しない「きよしこの夜」の演奏の1つにおいて、本手法による小節区切れ位置の推定にわずかに失敗した。これは、同一フレーズの総数による小節切り出し候補の絞り込みで誤って絞り込んでしまったものと思われる。

FINALE では、自動採譜の際に MIDI データのテンポおよびビート情報を利用するので、メトロノームを利用しない演奏について、FINALE はうまく推測できない。メトロノームを利用しない演奏について、特に本手法に良好な結果が得られた。

4.3 同一フレーズの抽出率

表 4.4に、メトロノームを使用した演奏について、楽曲毎に、同一フレーズの抽出率を調べて示す。

表 4.4: 同一フレーズの抽出率（メトロノーム使用）

曲目	FINALE	本手法
「アルプス 1 万尺」	47%	100%
「荒城の月」	53%	92%
「今日の日はさようなら」	33%	100%
「きよしこの夜」	44%	89%

（ 値は、それぞれ 3 つの演奏の平均）

表 4.5に、メトロノームを使用しない演奏について、楽曲毎に、同一フレーズの抽出率を調べて示す。

表 4.6に、演奏家による演奏について、演奏毎に、同一フレーズの抽出率を調べて示す。

研究者本人による演奏では、ほぼすべての同一フレーズを抽出できた。演奏家による演奏は、同一フレーズの抽出率が低下しているが、これは、演奏家が演奏に表情付けを行っているのが原因ではないかと考えられる。

表 4.5: 同一フレーズの抽出率（メトロノーム使用せず）

曲目	FINALE	本手法
「アルプス 1 万尺」	40%	93%
「荒城の月」	42%	100%
「今日の日はさようなら」	67%	100%
「きよしこの夜」	39%	94%

（ 値は、それぞれ 3 つの演奏の平均）

表 4.6: 同一フレーズの抽出率（演奏家の演奏）

曲目	FINALE	本手法
「きよしこの夜」演奏 1	39%	89%
「きよしこの夜」演奏 2	50%	78%
「荒城の月」	28%	81%

4.4 類似フレーズの抽出率

表 4.7に、メトロノームを使用した演奏について、楽曲毎に、類似フレーズの抽出率を調べて示す。

表 4.7: 類似フレーズの抽出率（メトロノーム使用）

曲目	FINALE	本手法
「アルプス 1 万尺」	14%	76%
「荒城の月」	22%	64%
「今日の日はさようなら」	15%	72%
「きよしこの夜」	37%	67%

（ 値は、それぞれ 3 つの演奏の平均）

表 4.8に、メトロノームを使用しない演奏について、楽曲毎に、類似フレーズの抽出率を調べて示す。

表 4.8: 類似フレーズの抽出率（メトロノーム使用せず）

曲目	FINALE	本手法
「アルプス 1 万尺」	12%	88%
「荒城の月」	14%	72%
「今日の日はさようなら」	31%	51%
「きよしこの夜」	11%	85%

（ 値は、それぞれ 3 つの演奏の平均）

表 4.9に、演奏家による演奏について、演奏毎に、類似フレーズの抽出率を調べて示す。

本手法では、ある程度の類似フレーズしか抽出できないことが分かった。研究者本人による演奏は、各類似フレーズの演奏にばらつきがあるが、演奏家による演奏は、それらと較べて、あまり発音パターンの誤差が少なかった。しかし、演奏家による演奏は、採譜の

表 4.9: 類似フレーズの抽出率（演奏家の演奏）

曲目	FINALE	本手法
「きよしこの夜」演奏 1	30%	85%
「きよしこの夜」演奏 2	26%	63%
「荒城の月」	19%	56%

ための演奏ではなく、テンポ変動を自由に行った演奏であり、それが原因で類似フレーズとして抽出されない部分が見られた。

4.5 自動採譜システムの改良

類似フレーズは、その発音パターンにばらつきがあるので、同一フレーズのようにうまく抽出できない。そこで、より類似度の強い新しい類似フレーズを定義して、自動採譜システムの改良を図る。

4.5.1 小節内のピッチの上下動に関する類似性の利用

小節内の演奏について、

- 構成音数とそれらのピッチのならびが等しい
- となり合う演奏音のピッチの上下動の様子が等しい
- 構成音の発音パターンが類似である

とき、それらは、より類似度の強い新しい類似フレーズとする。この新しい類似フレーズには、より大きな発音パターンの誤差を認めてやる。これにより、より多くの類似フレーズの抽出が期待される。

4.5.2 新しい類似フレーズの具体例

図 4.1に、新しく導入した類似フレーズの例を示す。この楽譜において、第 1 小節と第 3 小節が、また第 2 小節と第 4 小節が、それぞれ新しい類似フレーズの関係にある。

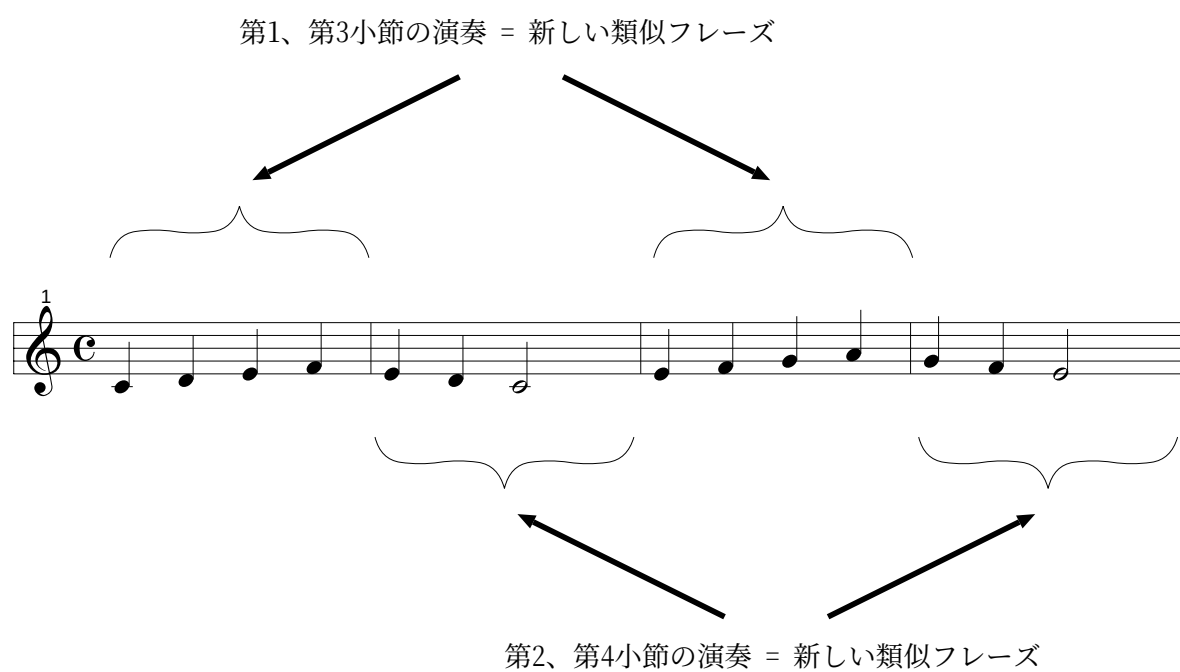


図 4.1: 新しい類似フレーズ

本論文の評価実験に用いている「アルプス 1 万尺」では、類似フレーズ c は、同一フレーズ C が 2 個と、その他のフレーズが 1 個の計 3 個で構成されているが、これら 3 個の類似フレーズはすべて新しく導入した類似フレーズの関係にある。同様にして、3 個で構成される類似フレーズ b は、そのうち、2 個が新しい類似フレーズの関係にある。新しい類似フレーズは、発音パターンの誤差を少し大きく持たせることができるので、従来の類似フレーズとして抽出されなかったものが、新たに抽出される可能性がある。

4.5.3 新しい類似フレーズ導入の効果

表 4.10 に、メトロノームを使用しない演奏を楽曲毎に 1 ずつ用意し、新しい類似フレーズを導入して類似フレーズの抽出を行った結果を示す。新しい類似フレーズの発音パターンの最大許容誤差を、6 として類似フレーズの抽出を行った。

表 4.10: 類似フレーズの抽出数（新しい類似フレーズ導入前／後）

曲目	新しい類似フレーズ数	類似フレーズの抽出数	
		導入前	導入後
「アルプス 1 万尺」	$b \times 2, c \times 3$	10	12
「荒城の月」	—	8	8
「今日の日はさようなら」	$c \times 2, d \times 2$	8	9
「きよしこの夜」	$b \times 2, c \times 2$	3	6

新しい類似フレーズの導入により、例えば、「アルプス 1 万尺」では従来の方法では 10 個の類似フレーズを抽出していたのに対し、この抽出方法では、12 個の類似フレーズを抽出することができた。同様に、「今日の日はさようなら」では、8 から 9 に、「きよしこの夜」では 3 から 6 へと類似フレーズの抽出率が上昇している。新しい類似フレーズを持たない「荒城の月」では、抽出率に変化はなかった。

第 5 章

結論

5.1 まとめ

本論文では、音楽構造に着目した自動採譜法を提案し、類似フレーズに着目した自動採譜システムについて、人間によるピアノ演奏を用いてその評価実験を行った。

類似フレーズに着目した自動採譜システムの評価実験では、題材として、類似フレーズを含む単純な曲を 4 曲選びだし、研究者本人が電子ピアノで演奏して、演奏データを収集した。システムに入力される人間による演奏は、採譜のための演奏ということを想定して、大きなテンポ変動を含まない演奏として、メトロノームに合わせた演奏、またはメトロノームを用いない、微妙なテンポ変動を含む演奏を用意した。

実験の結果、1 つの演奏に対し複数種類の小節切り出しの候補をつくり出すことが分かった。そして、それらの複数存在する候補それぞれの同一フレーズの総数を調べ、最もその総数の多い候補が正しい小節の区切り方を行っているものであることが分かった。

類似／同一フレーズの検査は、その構成音の発音パターンにある程度の誤差を認めて、パターンの類似なものを選び出す必要のあることが分かった。類似と認める誤差の大きさは、本実験では、4 とするのが適当であるという結果になった。それ以上の誤差を類似と認めてやると、本来類似フレーズでないものまで、ある類似フレーズのグループに含んでしまい、良い結果が得られなかった。

本手法で得られた楽譜を、演奏に用いたもとの楽譜と比較したところ、小節は、ほぼすべて正しい位置で区切られていることが分かった。また、同一フレーズを、ほぼすべて抽出できていることが分かった。しかし、類似フレーズはすべてを抽出することはできな

かった。

類似フレーズの可能性のあるフレーズのうち、音階の上下の様子が同じものには発音パターンの許容誤差を少し大きく認めてやることで、わずかではあるが、類似フレーズの抽出精度を高めることができた。

5.2 今後の課題

小節の切り出し位置や同一フレーズの抽出とは異なり、類似フレーズの抽出は、まだまだ満足できる精度で行われていないことが分かった。

また、1 拍目から始まらない演奏について、まずそのまま先頭の音を 1 拍目として楽譜データを作成し、MIDI データを再生成する際に、演奏の先頭に無発音時間を追加してやることで対応することが考えられる。

謝辞

本研究を行うにあたり、御指導および御鞭撻を戴いた北陸先端科学技術大学院大学 情報科学研究科 堀口進教授に深く感謝致します。

また、御指導、御助言を頂きました同研究科 阿部 亨助教授に深く感謝致します。

さらに、同研究科 示村悦二郎教授には、サブテーマで熱心にご指導を戴き、厚く御礼申し上げます。

同研究科 山森一人助手には、有益な御助言を頂くなど大変お世話になり、厚く感謝いたします。

そして、常日頃よりお世話になりました当研究室の皆様に心より感謝いたします。

参考文献

- [1] 平田圭二, 鈴木孝, 柏野邦夫, 特集「音楽情報処理」の編集にあたって, 情報処理, Vol. 35, No. 9, pp799–800, 1994.
- [2] 長嶋洋一, 音素材の生成, 情報処理, Vol. 35, No. 9, pp808–814, 1994.
- [3] Chowning, J. M., The Synthesis of Complex Audio Spectra by Means of Frequency Modulation, Journal of Audio Eng. Soc., 21(7), pp526–534, 1973.
- [4] 田口友康, 自動演奏と演奏モデル：演奏の芸術性をめざして, コンピュータと音楽 (bit 別冊), pp42–53, 1987, 共立出版, 音楽情報科学研究会 (編) .
- [5] 福岡俊之, 片寄晴弘, 井口征士, 音楽解釈システム MIS における自動演奏生成過程について, 情報処理学会 第 40 回 (平成 2 年度前期) 全国大会 講演論文集, pp1585–1586, 1990.
- [6] 玉城謙一, 片寄晴弘, 井口征士, 演奏ルールの抽出と再現, 情報処理学会 研究報告, 94(103), 13-16, 音楽情報科学 研究報告 No.8 於 沖縄芸大.
- [7] 玉城謙一, 片寄晴弘, 井口征士, 音楽表現の要因に関する一検討, 情報処理学会 第 48 回 (平成 6 年度前期) 全国大会 講演論文集, pp1.373–1.374.
- [8] 岡謙太郎, 楽譜印刷, コンピュータと音楽 (bit 別冊), pp112–122, 共立出版, 音楽情報科学研究会 (編) .
- [9] 大照完, 楽譜の自動認識, コンピュータと音楽 (bit 別冊), pp92–100, 共立出版, 音楽情報科学研究会 (編) .
- [10] 片寄晴弘, 今井正和, 井口征士, 感性情報抽出のための知的採譜システム, 人工知能学会 第 2 回 全国大会 論文集, pp585–588, 1989.

- [11] 片寄晴弘, 井口征士, 知的採譜システム, 人工知能学会誌, 5(1), pp59–66, 1990.
- [12] 柏野邦夫, 田中英彦, 計算機への音楽の入力—「音」の分離抽出の難しさ—, 情報処理, Vol. 35, No. 9, pp801–807, 1994.
- [13] 平賀譲, 音楽認知への計算機的アプローチとその課題, 情報処理, Vol. 35, No. 9, pp822–829, 1994.
- [14] Desain, P., & Honing, H., Quantization of Musical Time; A Connectionist Approach, Computer Music Journal, vol. 13, pp56–66, 1989.
- [15] Rosenthal, D., Intelligent rhythm tracking , NL-84-13, pp95–102, 1991.
- [16] Allen, P., & Dannenberg, R., Tracking Musical Beats in Real Time Proc. of the International Computer Music Conference, pp140–143, 1990.
- [17] Desain, P., & Honing, H., Foot-tapping: a brief introduction to beat induction, Proc. of the International Computer Music Conference, pp78–79, 1994.
- [18] 土屋景一, パターン認識を用いた音楽情報データベースの圧縮法, 情報処理学会研究報告, Vol. 96, No. 53, pp45–50, 1996.

研究業績

- [1] 「類似フレーズに着目した人間によるピアノ演奏の自動採譜」平成 9 年度電気関係学会北陸支部連合大会 11 月 2-3 日 金沢工業大学