

卒業論文

即興演奏のためのメロディーから 感性に合うコード譜の作成支援システム

Support system for creating chord scores that match your
sensibilities from melodies for improvisational performances.

富山県立大学 工学部 情報システム工学科

2020043 山内拓海

指導教員 António Oliveira Nzinga René 講師

提出年月: 令和6年(2024年)2月

目次

図一覧	ii
表一覧	iii
記号一覧	iv
第1章 はじめに	1
§ 1.1 本研究の背景	1
§ 1.2 本研究の目的	1
§ 1.3 本論文の概要	2
第2章 従来の研究	4
§ 2.1 メロディーを表現する既存のシステム	4
§ 2.2 ギターに関する専門知識	6
§ 2.3 リアルタイムでのコード譜作成の必要性	8
第3章 音楽分析のアプローチと印象の変化	12
§ 3.1 2つの楽譜の数値化による比較	12
§ 3.2 メロディーから音階のリアルタイム抽出	13
§ 3.3 コードの違いによる印象の変化	15
第4章 提案手法	19
§ 4.1 現存のコード譜と作成したコードの比較	19
§ 4.2 タブ情報の保存システム	20
§ 4.3 メロディーからコード譜作成システムの開発	22
第5章 数値実験並びに考察	27
§ 5.1 数値実験の概要	27
§ 5.2 実験結果と考察	27
第6章 おわりに	29
謝辞	30
参考文献	31

図一覧

2.1	音楽情報検索の例	5
2.2	デジタル音楽ワークステーションの例	5
2.3	FIMM の例	7
2.4	デジタル音楽ワークステーションの例	7
2.5	弦の説明	8
2.6	フレットの説明	8
2.7	音高推定の例	11
2.8	音高推定の例	11
3.1	MIDI でのノートナンバー	13
3.2	音階判定の様子	14
3.3	印象の違い	16
3.4	メジャーコードの構成	18
3.5	マイナーコードの構成	18
4.1	コードトーン	20
4.2	Amajor	20
4.3	B マイナーコードトーン	21
4.4	Bminor	21
4.5	B メジャーコードの情報	22
4.6	F メジャーコードの情報	22
4.7	システム全体の流れ	23
4.8	メジャーコード	25
4.9	マイナーコード	25

表一覧

5.1 アンケート項目	28
5.2 アンケート結果	28

記号一覧

以下に本論文において用いられる用語と記号の対応表を示す.

用語	記号
トニック機能	T
ドミナント機能	D
フーリエ変換の周波数成分	c_k
離散時間信号の値	f_n
フーリエ変換の複素指数項	$e^{-i \cdot 2\pi / N}$
N 個のサンプルに対する回転因子	w_N
偶数番目のサンプル	f_n^c
奇数番目のサンプル	f_n^o
$N/2$ 個のサンプルに対する回転因子	$w_{N/2}$
N 個のサンプルに対する回転因子	W_N

はじめに

§ 1.1 本研究の背景

音楽の創造的な表現は人類が持つ豊かな文化の一翼を担い、感情やアイデンティティの表現において重要な役割を果たしている。音楽家やアーティストは、耳コピや即興演奏のスキルを駆使して独自の音楽を生み出し、聴衆との共感を生み出している。しかし、これらのスキルを磨く過程は熟練を要し、時には限定的な情報やリソースに依存することがある。

従来の音楽理論や楽器の技術向上だけでなくデジタル技術の進化により、新たな音楽制作の手法やサポートツールが登場している。しかし、依然として楽曲の細部まで把握し自らのアイデアを具現化するための効果的な手段が求められている。例えば、複雑なコード進行や微細なフレーズを正確に理解することはアーティストが自分の音楽に深みを与える上で重要である [1]。

現代の音楽制作においてアーティストやミュージシャンは多岐にわたる音響情報を取り込み、それを元に独自の音楽表現を構築している。しかし、そのプロセスは手作業や経験に頼る部分が依然として存在し特に耳コピや即興演奏においては、正確かつ迅速な情報処理が求められている。この課題に対処するためにデジタル技術を駆使した新しい手法やツールの開発が期待されている [2]。

この背景から、本研究では音楽情報処理技術の応用によってアーティストの創造的表現を支援する手法の確立を目指す。具体的には、入力された音楽をリアルタイムで解析し、コード進行やメロディの特徴などを提示するシステムを構築する。これによってアーティストは自身の演奏内容を迅速に把握でき細部への気づきを高めることが期待できる。さらに、解析結果に基づいた音楽生成技術を応用することでアイデア拡張や自動伴奏などの新たな創作支援を実現できる可能性がある。こうした音楽制作支援システムは、アーティストの限界を超えた創造性向上に大きく貢献すると考えられる。新しい手法とツールの開発はアーティストたちがより豊かな音楽を生み出すための手助けとなると考えられる。

§ 1.2 本研究の目的

ギタリストが効果的に耳コピや即興演奏を行うための支援が求められている。バンドでの演奏において、エレキギターは他の楽器との調和を取りながら特有のサウンドを提供するためその技術的な要素や表現力の向上が不可欠である。

本研究の目的は、ギタリストが耳コピと即興演奏を効果的に行うための支援システムを構築することを目指している。エレキギターの音楽的な表現には音符情報だけでなく、アーティキュレーションやニュアンスなど、独特の要素が含まれている。これらを模倣し演奏者が自在に表現できるようにすることが重要である [3]。特にギターのコードには多様な種類が存在し、これらを正確に理解し再現することが挑戦的である。そのため、構築するシステムは異なるコードの特徴を認識し、演奏者に適切なサポートを提供できるようになることが期待されている [4]。

この研究の成果は、ギタリストが高度な音楽的なスキルを持たなくても耳コピや即興演奏をより容易に行えるようにすることに寄与することが期待されている。また、バンド演奏においてエレキギターの役割を担う演奏者にとって音楽制作の幅が広がり、表現力が向上する可能性がある [5]。

また、リアルタイムで音楽的情報を解析しギタリストの演奏をサポートするシステムの構築には音声処理技術が不可欠である。演奏中の音の長さ、メロディーの変化を瞬時に解析しそれをもとに楽曲全体のコンテキストを理解することが求められる [6]。このシステムは、演奏者が歌ったメロディーを捉え、それを使用されたスケールやコード進行と結びつけることが重要である。同時に、リアルタイムでリズムやテンポの変化を正確に把握し演奏者に最適な情報を提供している。FFT を用いた音響信号処理の技術を活用して演奏者に支援を提供する。また、使いやすいインターフェースを備え演奏者が自由に表現できるようにする。

高度な音楽的なスキルがなくとも演奏ができるようわかりやすい指板に抑える場所をマークしてあるコード譜を生成し出力を行う。この音楽情報解析に成功すれば演奏者は即興演奏中もシステムからの確かな情報を得ることができ、アイディアの迅速な発展や演奏のクオリティ向上が期待される。この支援システムではギタリストにとって創造性と表現力を高める重要なツールとなることを目指す。

§ 1.3 本論文の概要

本論文は次のように構成される。

第1章 本研究の背景と目的について説明する。音楽分野における即興演奏や耳コピの重要性について述べる。目的は背景で挙げた課題に対して、リアルタイムでの音高推定を用いて演奏支援システムの提案について述べる。

第2章 音楽分野における従来研究についてまとめ、リアルタイムでの tab 譜のについて述べる。また、ギターについての名称や専門知識についての説明を行う。

第3章 本研究の提案手法に用いる音楽理論の一部について解説する。また、提案手法のシステム部分に用いる音階の抽出方法について解説する。

第4章 本研究の提案手法について述べる。

第5章 実際にシステムを使用してもらい評価を行った。ギターを演奏してもらい有効性の検証を行った。

第6章 本論文における前章までの内容をまとめつつ，本研究で実現できたことと今後の展望について述べる．

従来の研究

§ 2.1 メロディーを表現する既存のシステム

音楽の生成と合成

音楽のメロディー分析は多岐にわたる音楽関連の分野で重要な役割を果たしており、その重要性はますます高まっている。この分野では音楽理論、楽曲制作、音楽情報検索、音楽教育などが相互に補完し合い進化し続けている。音楽理論の文脈ではメロディーは楽曲の中で聴衆に最も直接訴えかける要素の一つである。メロディーの分析は異なる時代やジャンルの楽曲におけるメロディーの構造や進行を理解する手段である。音楽理論家や研究者はメロディーにおける音程やリズムのパターン、そしてハーモニーとの関係性を分析し、作曲家がどのように聴衆の感情に訴えかけるかを解明している。

楽曲制作の現場ではデジタル音楽ワークステーション (DAW) や楽譜ソフトウェアが広く活用されている。これらのソフトウェアはメロディー分析を通じて作曲家やプロデューサーに対して革新的な手段を提供している。メロディーの自動生成や編集、アレンジの支援など、様々な機能が導入されアーティストたちはこれらのツールを駆使して創造的なプロセスを進めている。音楽情報検索システムは大規模な音楽データベース内で楽曲を見つける際にメロディー分析を駆使している。これにより、ユーザーは特定のメロディーを持つ楽曲を容易に見つけたり関連する楽曲を発見したりできる。アルゴリズムや機械学習の進歩により、高度なメロディー認識が可能となりユーザーエクスペリエンスが向上している。

音楽教育分野ではメロディー分析が学習者が楽曲を理解し、演奏するための基盤となっている。教育用ソフトウェアや教材は学習者が楽曲内のメロディーの構造や音楽理論を効果的に理解できるよう工夫されている。また、メロディーの分析を通じて学習者が創造的な表現力を養う手段ともなっている。最近ではVRや拡張現実を用いた新しい音楽教育プラットフォームも登場している。これらはメロディーの視覚化やインタラクティブな操作を可能にし、より能動的で直感的な学習を実現している。将来的には脳信号に基づいた学習状況の把握や、AI アバターによるパーソナライズされた学習支援など、教育効果を飛躍的に高める革新的技術が登場することが期待されている。

音楽情報検索とデジタル音楽ワークステーション

音楽のメロディー分析は音楽理論や楽曲制作、音楽情報検索など、音楽研究のさまざまな分野で重要な役割を果たしている。これに関連して多くの既存のシステムが開発および研究されている。既存のシステムの一例として、音楽情報検索分野におけるシステムが挙げ

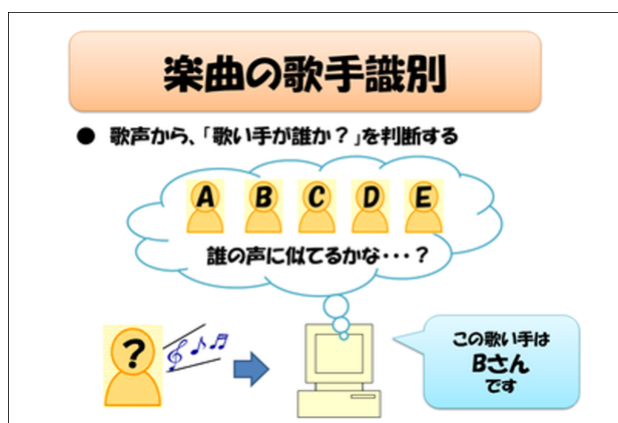


図 2.1: 音楽情報検索の例



図 2.2: デジタル音楽ワークステーションの例

られる。これらのシステムは楽曲のメロディーラインを自動的に抽出し、音楽データベース内で楽曲を検索したり、ジャンルを分類したりするのに使用されている。最近では、機械学習を用いた手法が取り入れられより高度なメロディー認識や楽曲検索が可能となっている。さらに、音楽制作の分野では、デジタル音楽ワークステーションソフトウェアや楽譜ソフトウェアがメロディー分析のために使用されている。これらのシステムは楽曲制作プロセスにおいてメロディーの作成、編集、アレンジに必要な機能を提供し、ミュージシャンや作曲家に創造的なツールを提供している。最近のシステムでは機械学習を用いてユーザーの創作スタイルを学習し、新しいメロディーやコード進行の提案を行うなどインタラクティブなメロディー制作を可能にしている。

また、音楽教育の分野でもメロディー分析が重要である。eラーニングシステムやゲーム形式の教育アプリなど、最新の教育技術がメロディーの認識や反復練習のために活用されている。これにより個別化され効果的な学習が実現している。

最後に、音楽制作プロセスにおいてメロディー分析はアーティストやプロデューサーにとっても重要である。メロディーラインの解析は楽曲の構造や楽器アレンジに関する洞察を提供し、音楽の創造的なプロセスを向上させる。最近のプロダクションシステムではメロディーやコード進行の複雑性や単調性などの指標を可視化しており、楽曲制作者の意思決定を支援している。これらのように、メロディーを分析する既存のシステムは音楽研究、制作、教育の各分野で多くの利点をもたらすこれらのシステムの進化は音楽の理解や制作プロセスの向上に寄与している。今後も最新のテクノロジーを取り入れながらメロディー分析システムの更なる高度化が期待されている。

AIとメロディー生成

近年、人工知能 (AI) の進化が音楽制作とメロディー生成の分野において注目を集めている。AIは機械学習やニューラルネットワークを活用し、既存のメロディーや楽曲から学習し、新しいメロディーを生成する能力を持っている。音楽制作において、AIはアーティストやプロデューサーに革新的な手段を提供している。AIモデルは大量の楽曲データを分析し、異なるジャンルやスタイルに基づいてメロディーを生成することができる。これにより、創

造的なプロセスが効率的になり、新しい音楽のアイデアが生み出される可能性が広がっている。

一方で、AIのメロディー生成は音楽理論や感性の側面においても議論的となっている。アートの創造性や感情の表現において、AIが本当に人間と同様の深さを持つことができるのか、といった問いかけがなされている。これらの課題は、AIとメロディー生成の関係性において新たな研究やディスカッションを呼び起こしている。さらに、AIを活用した音楽情報検索システムも進化している。AIは高度なメロディー認識技術を駆使して、ユーザーが特定のメロディーを持つ楽曲を見つけたり、関連する楽曲を推薦したりする際に役立っている。これにより、音楽のディスカバリー体験が向上し、ユーザーが新しい音楽に出会いやすくなっている。加えて、AIを用いたインタラクティブなメロディー編集ツールも開発されつつある。これらはユーザーのメロディー入力に対して、拍子や調性などの音楽理論的妥当性を判断し、適切な推薦を行う。拍子感や旋律のつながりを保ちつつ、ユーザーの創造性を高める編集が可能となる。

総じて、AIとメロディー生成の結びつきは、音楽の創造的な領域や技術の進歩に新たな可能性をもたらしている。アートとテクノロジーが融合する中で、今後ますます興味深い展望が広がると考えられる。特に、人間の創造性とAIの能力を最適に組み合わせた新しい創作プロセスの確立が期待されている。この点に関する研究がさらに加速することで、音楽制作におけるパラダイムシフトが起きるかもしれない。

AIによるメロディー生成の一例を紹介する。

FIMMIGRM FIMMIGRMはAIによるメロディー生成を基盤としたオリジナル楽曲制作サービスだ。FIMMIGRM™は、従来のBGM生成サービスとは異なりユーザーが独自のメロディーを作成し編集、ボーカルを追加することが可能だ。FIMMIGRMの最大の特徴はAIが生成した楽曲がクリエイターやプロフェッショナルにとって利用しやすいことである。ユーザーはGENERS, GENERATE, PRO-ARRANGEDの3つの異なる方法で楽曲を作成できその過程で煩雑な手続きや専門的な知識を必要としていない。

また、FIMMIGRMはその高品質なアウトプットと無限の生成パターン数で注目されている。クリエイターは、簡単にキャッチーでハイクオリティなオリジナル楽曲を生成できインフルエンサーや動画クリエイターなど幅広い分野で利用されている。

このように、FIMMIGRMはAIによるメロディー生成技術の革新的な応用例であり音楽制作における新たな可能性を提示している。

§ 2.2 ギターに関する専門知識

ギターに関する専門知識において、弦とフレットについての基本的な理解は非常に重要である。ギターを演奏するためには、これらの要素について正確に理解しておく必要がある。ギターは、その特有の音色と多彩な奏法で知られる楽器である。ギターの演奏を始める際に最初に覚えるべき重要な要素の一つが、弦とフレットの呼び方だ。これらの用語を理解することは、ギター演奏の基礎知識として不可欠である。

弦 (Strings)

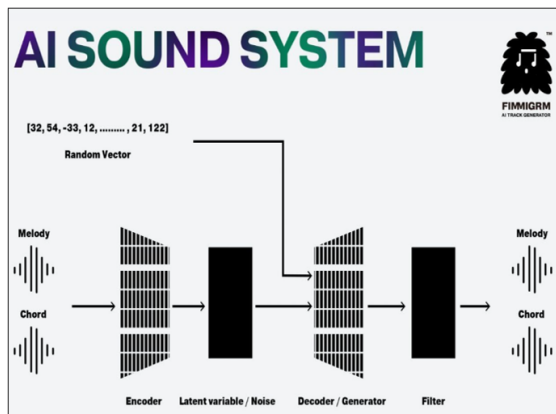


図 2.3: FIMM の例



図 2.4: デジタル音楽ワークステーションの例

ギターの弦は通常 6 本あり，一般的なチューニングは上から順に E（ミ），B（シ），G（ソ），D（レ），A（ラ），E（ミ）である．これらの弦は低音から高音まで順に配置されている．弦は通常，鋼弦やナイロン弦などの異なる材質で作られており，その材質によって音色が変わる．一般的なアコースティックギターは鋼弦を使用し，一方でクラシックギターは通常ナイロン弦を好んで使用している．弦の選択は演奏する音楽のスタイルや好みによって異なりますが，それぞれの材質が独自の音質を生み出している．

フレット (Frets)

コード譜はギターの演奏において使用される記譜法で，弦とフレットの位置を直感的に示している．横のラインはギターの弦を，縦のラインはフレットを表し，各弦に対して特定のフレットで指を押さえるかどうか数字や記号で示される．例えば，「C」コードは，3 番目の弦を 1 フレットで押さえ，2 番目の弦を 2 フレットで押さえ，1 番目の弦を 1 フレットで押さえることで演奏されている．コード譜を理解することで，演奏者は指定されたフレットと弦を押さえることで特定のコードを演奏し，楽曲を正確に表現することができる．

このようなコード譜の理解はギター演奏の基本であり，初心者から上級者まで，演奏者が楽曲を正確に再現するために欠かせず，練習と経験を積むことで，コード譜を見て即座に正確な指の動きをする技術が向上し，より高度な演奏が可能になる．

ギターのコード (Chords)

ギターのコードは複数の弦を同時に押さえて演奏することで構成されている．例えば C メジャーコード，G マイナーコード，D セブンスコードなどがあり，これらのコードの演奏には指の正確な位置や弦の適切な押さえ方が必要だ．練習を重ねることでコードの移動や切り替えがスムーズになり，より豊かな音楽表現が可能となる．ギターのコードは，指板上の特定のポジションで特定の弦を特定のフレットで押さえることで構成されている．例えば，C メジャーコードは 1 番目の弦を 1 フレット，2 番目の弦をオープン，3 番目の弦を 2 フレットで押さえることで演奏される．同様に，他のコードも特定の指の配置によって演奏される．

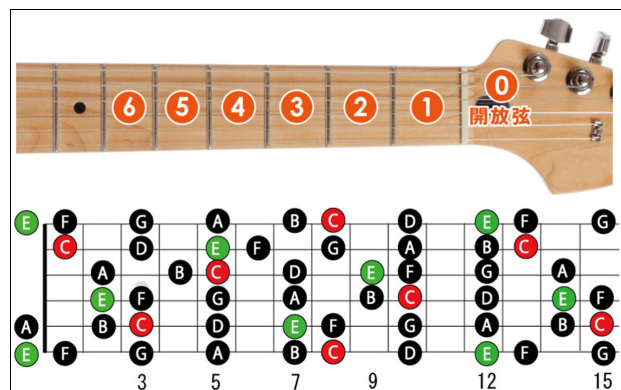


図 2.6: フレットの説明

弦の押さえ方とフレットの使い方

ギターを演奏する際、左手で弦を押さえ、右手で弾くことによって音が出ます。弦がフレット上で押さえられることで、弦の有効な長さが変わり、異なる音が発生する。フレット上で指を使って押さえることで、異なるコードや音階が演奏できる。ギターの指板上には様々なポジションがあり、これによって異なる音楽のフレーズを演奏できる。

指板上の音階

ギターのフレット上での音階の位置は、演奏者にとって理解が必要な重要な要素である。フレット上の音階は、特定の弦とフレットの組み合わせによって生成される音の配列で、これによってメロディやソロ演奏を構築することが可能となる。

各弦は低音から高音まで順に配置され、フレット上の位置によって音の高さが変化する。通常、1フレットから順に1, 2, 3,...と番号が振られ、同時に異なる弦上で同じフレットを押さえることで和音を形成できる。例えば、1弦（高音側のE弦）の3フレットで押さえた場合、それはG音となる。同様に、2弦の3フレットで押さえるとB音となる。これらの音を順に組み合わせることで、音階やコードを演奏できる。

また、ギターフレット上の音階は、パターンとして覚えることができる。これは、特定の音階（メジャー、マイナーなど）に基づいて、一定のフレットと弦上の指の配置が決まっているパターンである。これにより、演奏者は効果的に異なる音階をフレット上で表現し、創造的な演奏を構築することができる。練習と経験を通じて演奏者はフレット上の音階の位置を覚え、それを活かして表現豊かな演奏を実現することができる。

§ 2.3 リアルタイムでのコード譜作成の必要性

音楽の世界において、ギター演奏者がリアルタイムでコード譜を作成する必要性は多岐にわたる。演奏者が即興演奏を行う際、感情やインスピレーションをその場で表現する必

要がある。このとき、リアルタイムでのコード譜作成は、アイディアやメロディを整理し、他のミュージシャンとのセッションを促進している。同様に、セッション演奏においては、楽曲の構造や進行を共有することで、バンドやアンサンブル全体の調和が生まれる。

また、アレンジメントの柔軟性も重要な要素となる。楽曲のアレンジをリアルタイムで変更し、演奏者がその場でコード進行を調整することで、新しい要素を導入することができる。これにより、演奏の臨機応変な変化や表現力の向上が可能となる。教育面でも、リアルタイムでのコード譜作成は重要である。レッスンやワークショップにおいて、講師が生徒に即座にコード進行やメロディを示すことで、生徒は迅速に実践的な演奏スキルを身につけることができる。

これらの理由から、リアルタイムでのコード譜作成はギター演奏者やミュージシャンにとって欠かせないスキルであり、テクノロジーを駆使したツールやアプリケーションがますます注目されている。即興演奏やセッション、創造的な制作活動において、リアルタイムでのコード譜作成が音楽の豊かな表現力を支え、新たな音楽体験を生み出す重要な要素となっている。

そのため、これらに必要な音高推定の手法についていくつか紹介する。

自己相関関数を用いたピッチ検出

スペクトル解析は、音声信号を周波数成分に分解する手法である。この手法はピッチ検出において、主要な周波数成分を抽出し、その周期を用いて音の高さを推定する。高精度かつ幅広い楽器や声に対応する特性を持っている。

音声信号をスペクトル解析する過程では、信号がどの周波数成分から構成されているかを把握している。ピッチ検出では、その中で主要な周波数成分を特定し、その周期に基づいて音の高さを推定する。これにより、楽曲や発声から得られる音の情報を把握し、高度な音楽解析や音声処理が可能となる。この手法の利点は高い精度でのピッチ検出が可能であり、異なる楽器や声の音高変化にも対応できることである。音楽の表現や声の特徴において、主要な周波数成分がピッチ情報を提供しているため、スペクトル解析は広く用いられている。

スペクトル解析を用いたピッチ検出は、音楽理論や信号処理において基本的な手法の一つであり、その高精度な解析は音楽制作や音声処理の分野で重要な役割を果たしている。この手法を用いることで、音楽の理解や音声の変換、効果的な音響処理などが実現され、多岐にわたる応用が可能である。さらに、スペクトル解析に基づいて音声信号から抽出した周波数と振幅の情報は、単にピッチ検出のみならず、楽器の識別や音色の分析など、高度な音楽情報処理にも大いに活用されている。例えば、楽器ごとの固有の倍音構造を分析することで、演奏されている楽器の種類を機械的に認識したり、混在した音声の分離を行う音源分離が可能だ。これは自動採譜や音楽情報検索などの応用に大きな影響を与えている。加えて、スペクトル包絡やスペクトルの時間変化も抽出可能で、これに基づいて残響や音の立ち上がり特性といった音響環境の情報が得られる。これらは楽器の音色制御や音場再現の研究に大いに貢献している重要な知見といえる。

深層学習に基づく音高推定

深層学習アーキテクチャ、特にリカレントニューラルネットワーク (RNN) や長短期記憶 (LSTM) などを利用したモデルが音高推定に幅広く活用されている。これらのモデルは学習データから音高のパターンを学び、新しい音声入力に対して高精度な音高推定を行う。

RNN や LSTM は時間的な情報を扱うのに優れており、音楽データの連続性やリズムなど、時間的な依存関係を学習することができる。音高推定においては、楽曲内での音の変化やパターンを把握することが重要であり、深層学習モデルはその点で優れている。深層学習の利点の一つは、複雑な音楽データにも適用できることである。異なる楽器や複雑な和音など、多様な音楽要素を学習し、高い汎用性を持って音高推定を行う。これにより、ジャンルや楽器の異なるさまざまな音楽に対しても適切な推定が可能である。

さらに、深層学習モデルは大規模なデータから特徴とパターンを自動で抽出することができるため、人手による特徴エンジニアリングが不要となる。これによって、膨大な量の音楽データを効率的に学習し、精度の高い音高推定器を構築できる。データさえ揃えば、モデルの汎化性能が大きく向上するため、未知の楽曲に対する頑健性が高まる。加えて、深層学習モデルは損失関数やネットワーク構造などのパラメータチューニングによって最適化することが可能である。これらのハイパーパラメータを調整することで、対象とする音楽ジャンルやデータセットの特性に応じたカスタマイズが実現できる。様々な実験を重ねてモデルの性能を最大化することが期待でき、応用範囲の拡大につながっている。

これらの利点から、深層学習に基づく手法は今後も音高推定をはじめとした音楽情報処理の分野で大きな影響力を持ち続けると考えられる。その可能性は着実に拡大しつつある。

スペクトル解析に基づくピッチ検出

スペクトル解析は音声信号を周波数成分に分解する手法であり、ピッチ検出ではこれを利用して音の高さを推定する。この手法は高精度であり、さまざまな楽器や声に対応することができる。

音声信号は複雑な波形から構成されていますが、スペクトル解析を用いることでその周波数成分を明示的に抽出することが可能である。ピッチ検出では、主要な周波数成分を抽出し、その周期を用いて音の高さを推定する。これにより、どの周波数が支配的であるかを特定し、それに基づいて正確な音高情報を得ることができる。この手法の特長は、高い精度で音の高さを推定できることである。また、異なる楽器や声の特性にも適応可能であるため、多岐にわたる音楽ジャンルや声の種類に対して柔軟に利用することができる。楽器ごとの音色や声の個性を考慮した音高推定が可能であり、これは音楽制作や音声処理において重要な要素である。

スペクトル解析とピッチ検出の組み合わせは、音響信号処理において幅広い応用がある。例えば、楽曲のキーの特定、ボーカルのトラッキング、音声の変換などに利用され、音響技術の向上に寄与している。これらの手法は高度な技術を要するものの、その結果として得られる精緻な音楽や音声の制御が可能である。最近ではディープラーニングを組み合わせた手法も登場しており、複雑で多様な音楽シーンに対して汎用的に適用可能なピッチ検出器が実現しつつある。大量の音楽データから特徴と周波数構造を自動で獲得することが可能となっており、さらなる精度向上が期待されている。

FFT を用いた音高推定

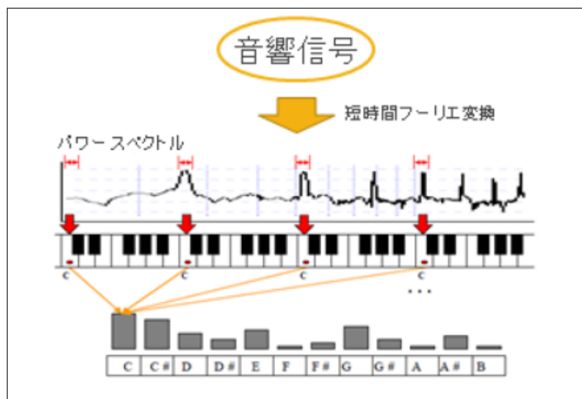


図 2.7: 音高推定の例

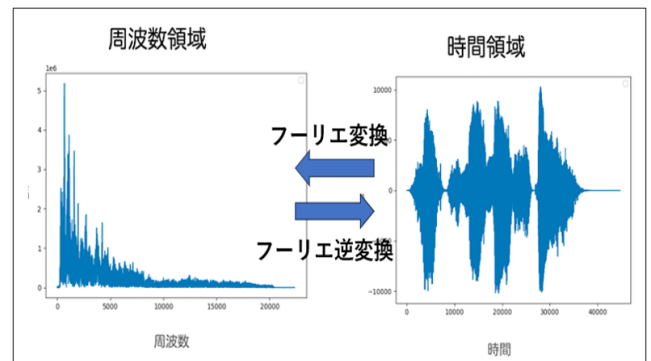


図 2.8: 音高推定の例

音高推定には様々な手法があり、その一つがFFT（高速フーリエ変換）を用いた手法だ。この手法では、音声信号を短い時間フレームに分割し、各フレームにFFTを適用して周波数成分を取得している。これにより、高い分解能で音の高さを推定することが可能である。得られた周波数成分を解析し、主要な周波数やピークを検出して各フレームごとの音の高さを推定し、連続的に統合して滑らかな音高曲線を形成している。

この手法は複雑な楽曲や細かな音高変化を捉えるのに優れていますが、計算コストが増えるため、リアルタイムの処理や大量のデータに対する適用には計算リソースが必要である。一方で、高度な音楽解析に広く利用され、音声信号の周波数成分を詳細に分析することで精密な音高推定が実現されている。

音楽分析のアプローチと印象の変化

§ 3.1 2つの楽譜の数値化による比較

まず最初に、オクターブと MIDI について説明する。

オクターブ

音楽における「オクターブ」と「半音」の関係について説明する。音楽理論に精通していない読者にも理解しやすいよう、平易な言葉を用いて丁寧に記述する。音の周波数の2倍の整数比関係にある2つの音を、1オクターブの間隔がある音と定義する。例えば、ある音の周波数が f Hz であるとき、その音の完全1オクターブ上の音の周波数は $2f$ Hz となる。このとき2つの音の比は $1:2$ であり、波長の比は $1:1/2$ の関係が成り立つ。逆に1オクターブ下の音の周波数は $f/2$ Hz となる関係が成り立つ。次に、ある音から半音上へ移った場合、周波数が約6%増加することを示す。例えば、ある音の周波数が 440 Hz である場合、その半音上の音の周波数は $440 \times 2^{(1/12)} \div 466$ Hz と計算できる。ここで2の $1/12$ 乗は、12平均律における半音の周波数比であり、約 1.05946 である。同様に半音下の音は、440Hz の約6%減で約 414Hz の周波数となる関係が成り立つ。1オクターブには12個の半音が均等に配置されているため、半音は音階を表す最小単位であると考えられる。

上記のオクターブと半音の関係性を用いて、基準となるある音から周波数が整数倍となる任意の音を導出することができる。ここで基準の音をピアノの基本オクターブにある C2(130.81 Hz) とし、周波数が2倍(261.63 Hz)の関係にある C3、3倍(392.44 Hz)の G3、4倍(523.25 Hz)の C4などを順次算出していく。この結果は、十二平均律に基づく音楽の標準的な周波数体系を説明し、音程と周波数の関係性を具体的かつ詳細に示す上で有用であると考えられる。

MIDI

MIDI ノート番号は音楽業界で一般的に使用される音階の数値表現方法である。C4を基準にして、音階ごとに整数が割り当てられる。これを図 3.1 に示す。例えば、C4は MIDI ノート番号 60 に対応し、それ以降の音階は相対的に割り当てられる。この手法により、異なる音階をデジタルで正確に表現し、制御することが可能となる。MIDI ノート番号の利用は、音楽制作において楽曲の作成、演奏、録音、編集などにおいて非常に有益であり、音楽理論と数学的な計算を組み合わせることで高度な操作も可能である [4]。

この手法を応用することで、異なる楽譜の数値データを得ることができ、これを比較することで楽曲の違いや類似性を定量的に評価することができる。例えば、2つの楽譜の MIDI

60	C4	C3	261.6	48	C3	C2	130.8
59	B3	B2	246.9	47	B2	B1	123.5
58	A#3	A#2	233.1	46	A#2	A#1	116.5
57	A3	A2	220.0	45	A2	A1	110.0
56	G#3	G#2	207.7	44	G#2	G#1	103.8
55	G3	G2	196.0	43	G2	G1	98.0
54	F#3	F#2	185.0	42	F#2	F#1	92.5
53	F3	F2	174.6	41	F2	F1	87.3
52	E3	E2	164.8	40	E2	E1	82.4
51	D#3	D#2	155.6	39	D#2	D#1	77.8
50	D3	D2	146.8	38	D2	D1	73.4
49	C#3	C#2	138.6	37	C#2	C#1	69.3

図 3.1: MIDI でのノートナンバー

ノート番号を比較することで、音の高さやパターンの変化を明らかにし、作曲家のスタイルや楽曲構造の違いを理解する手段となる。これにより、楽曲の数値化が音楽分析に新たな視点を提供し、音楽の比較や理解において貴重なツールとなる。

楽譜を数値データに変換する際、音高と音価の両方を考慮することが重要である。音高の数値化においては、MIDI ノート番号を利用することで楽譜上の音の高さを整数で表現することができる。また、音価の数値化では、楽譜上の音符の長さを時間の単位で表現することが必要であり、これによりリズムやテンポの情報も含まれる。

比較手法としては、数値データを用いて類似性尺度を計算することが挙げられる。例えば、ユークリッド距離やコサイン類似度を利用して、2つの楽譜の類似性を評価することができる。これにより、作曲家や楽曲の異なる側面に焦点を当て、音楽の構造や特徴を数値的に捉えることが可能となる。

この手法の応用は多岐にわたり、音楽のジャンルや時代、作曲家の違いを明らかにするだけでなく、機械学習アルゴリズムを用いて楽曲の自動分類や類似楽曲の検索にも貢献することが期待される。楽曲の数値データをもとにした比較手法は、音楽理論とデジタル技術の融合により、音楽研究や楽曲制作の新たな可能性を切り拓いている。

§ 3.2 メロディーから音階のリアルタイム抽出

近年、音楽と情報技術の積極的な融合が進み、デジタルサウンド処理や人工知能を活用した様々なミュージックテクノロジーアプリケーションが登場している。そうした最新の音楽技術の一つとして、リアルタイムにメロディーから音階を抽出するシステムが注目されている。このようなリアルタイム音階抽出システムの基本的な仕組みと特徴について詳述する。

メロディーとは、楽曲において主旋律を構成する音の並びであり、楽曲の骨格をなす最も基本的な要素である。一方で音階とは、そのメロディー線の中で使用されている音の高さ関係を表すものである。従来のアコースティック楽器の演奏や作曲制作では、演奏者や作曲者が耳でメロディーの音高を聴き取り、使用されている音階を理解する能力が必要だっ

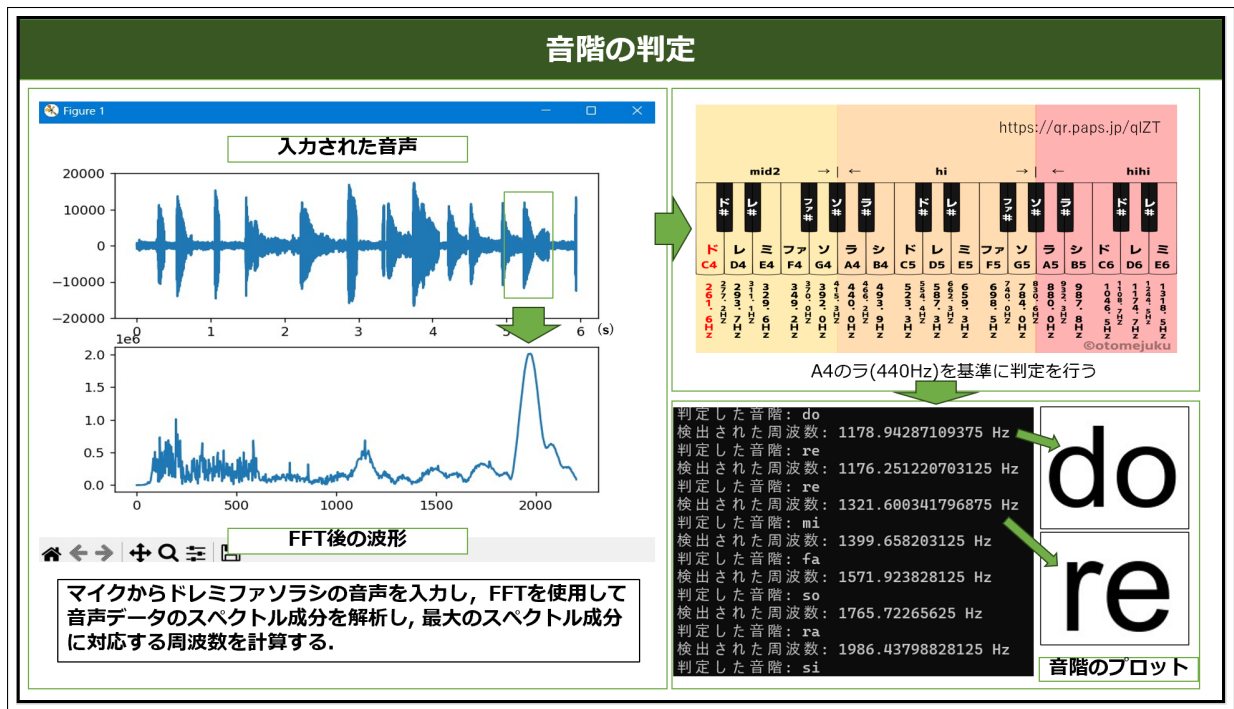


図 3.2: 音階判定の様子

た。しかし、コンピュータを使ってメロディーからリアルタイムに音階を自動抽出できるようになれば、音楽の制作や演奏のプロセスがより直感的かつ能率的になると共に、新たな音楽表現の方向性が拓かれることが期待されている。

提案するリアルタイム音階抽出システムは、マイク又はラインレベルの音声入力から取得した歌唱や楽器演奏のオーディオデータを処理対象とし、その中からメロディー部分をデジタル信号処理技術を用いて抽出した後、抽出されたメロディーデータから音楽理論に基づいてリアルタイムに音階を判定し、その情報をユーザーにフィードバックするものである。以下に、本システムの主要な構成要素と処理手順の流れを示す。

- ・オーディオ入力データ取得:P マイク用し、マイク等からリアルタイムに音声データを取得。
- ・メロディー抽出: 高速フーリエ変換の手法を用いて入力音声データからメロディーを抽出。
- ・音階解析: 抽出されたメロディーの周波数スペクトルを解析し、音階を推定。
- ・結果出力: 判定されたリアルタイムの音階データをユーザーにフィードバック。

< 高速フーリエ変換 (FFT)

$$\begin{aligned}
 c_k &= \sum_{n=0}^{N-1} f_n \cdot e^{-i2\pi nk/N} \\
 &= \sum_{n=0}^{N/(2-1)} f_{(2n)} \cdot e^{-i2\pi nk/(N/2)} + e^{-i \cdot 2\pi nk/N} \sum_{n=0}^{N/(2-1)} f_{(2n+1)} \cdot e^{-i2\pi nk/(N/2)}
 \end{aligned} \tag{3.1}$$

複素数空間の回転単位 $w_N = e^{-i \cdot 2\pi/N}$ とし、

偶数添字と奇数添字を $f_n^c = f_{2n}$ と $f_n^o = f_{2n+1}$ とおくと、

$$c_k = \sum_{n=0}^{N-1} f_n w_N^{kn} = \sum_{n=0}^{N/(2-1)} f_n^c \cdot w_{N/2}^{kn} + W_N^k \sum_{n=0}^{N/(2-1)} f_n^o \cdot w_{N/2}^{kn} \quad (3.2)$$

FFT によって音声データは周波数スペクトル成分に変換され、その中からメロディーを担う主要な周波数領域を抽出する。次いで、この抽出されたメロディデータの周波数パターンを解析し、現在の音高から推定される音階ラベルを音楽理論に基づいてリアルタイムに割り当てる。以上のように、本システムを用いることでユーザーは自身の歌唱や楽器演奏のメロディーから即座に現在の音階を取得することができる。

このようなリアルタイム音階抽出システムは、様々なマルチメディア分野やエンターテインメント分野における音楽コンテンツ制作現場での応用が期待されている。音楽制作者や演奏者が本システムからリアルタイムにフィードバックされる音階情報を利用することで、作曲や即興セッションをより直感的かつ能率的に進めることが可能になる。例えば、ギタリストが本システムを使用しながら新しいリフの即興演奏を行う場合を考える。ギターのフレット位置に対応した現在の音階データが表示されるので、次に演奏すべき音程の選択が格段に容易になる。

メロディーデータからリアルタイムに音階を抽出するシステムの基本概念と主要機能について述べた。このシステムは、新しい音楽の制作や演奏スタイルを切り拓く可能性を秘めていると考えられる。音楽家がより直感的かつ柔軟に楽曲制作を進められるようになる一方、新たなクリエイティブな表現が生まれることも期待されている。今後、本システムの高度化や様々な応用が進むことで、デジタルミュージック技術のさらなる発展につながることが予想される。

§ 3.3 コードの違いによる印象の変化

音楽において和音は異なる高さの音が3つ以上組み合わせられたもので、その構成や構成音の間隔によって印象が形成される。和音の感覚的な要素として構成音間の半音の数や和音の安定性が挙げられる。

感覚的な要素として構成音間の半音の数や和音の安定性が挙げられる。音楽理論ではこれらの感覚を半音の数やファンクションに基づいて理論化しており、例えば半音メジャーコードやマイナーコードが「明るい-暗い」を表現し、ドミナントモーションが「安定-不安定」を形成している。

和音の感覚は一般的に「明るい-暗い」や「安定-不安定」など共有されるものがある。これらの感覚は音楽理論で「明るい」を半音メジャーコードやマイナーコード、「安定」をドミナントモーションなどで表現しています。しかし、地域や文化によってはこれらの感覚が異なる。地中海沿岸では、特に短調の曲においてマイナー・メジャー・セブンスを主音とする感覚があるとされている。

地域や文化により、和音の印象が異なり、例えば、地中海沿岸では特有の感覚が報告さ

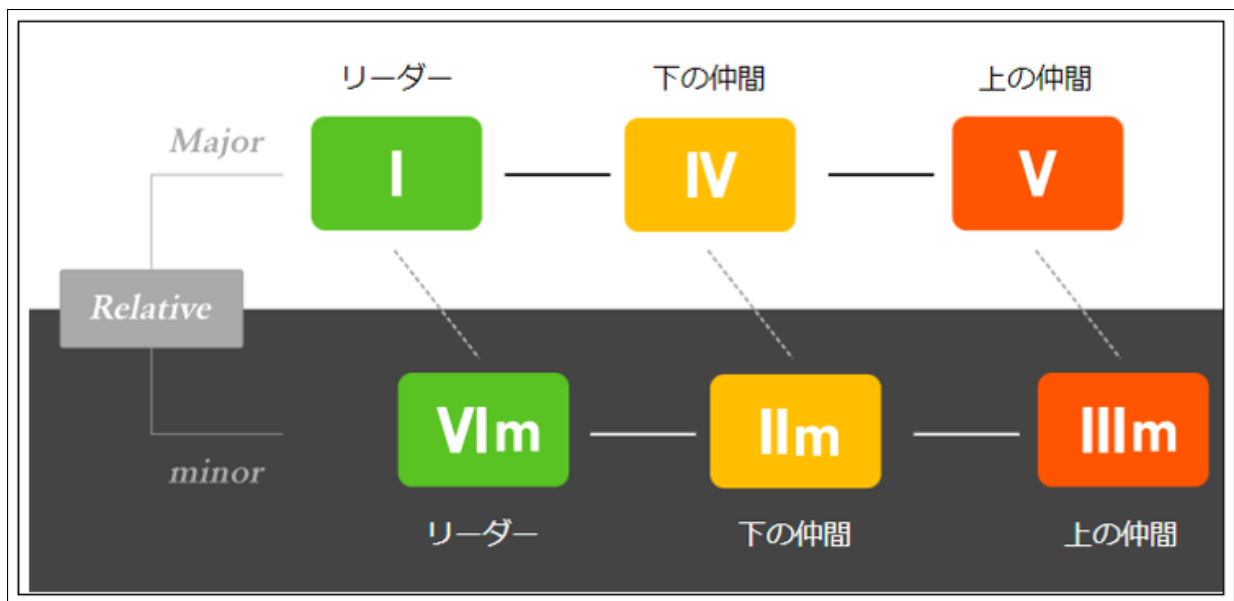


図 3.3: 印象の違い

れており、これらの差異を理解することで地域ごとの音楽文化や感性の違いが明らかになる。異なる文化や地域の感性に対する理解は、音楽の普遍性と地域差のバランスを考える上で重要である。

和音は楽曲において進行を導く要素として重要である。特定の和音の組み合わせやファンクションが、楽曲における安定性や不安定性を形成している。例えば、ドミナントからトニックへの進行は楽曲における解決の瞬間を生み出し、聴衆に感動を与えることがある。

和音の印象が聴衆に与える影響や異なる和音の組み合わせが感性にどう訴えかけるのかについても考察する。明るい印象を与えるメジャーコードと暗い印象を持つマイナーコードの対比や、和音の高低の配置が聴衆に与える心理的な影響について深く掘り下げる。また、和音の進行が楽曲において期待感や緊張感をどのように生み出すかを具体的な楽曲例を挙げて解説する。

音楽の和音はその構造や構成によって印象が変わり、地域や文化による感性の違いを理解し、音楽理論を通じて和音の印象を解明することで、楽曲制作や音楽理解に新たな視点が得られることが示される。和音の理解を通じて、音楽が人々の感性や文化とどのように対話しているのかを明らかにすることができる。

和音の構成音の間隔や進行が生み出す印象について、異なる地域や文化圏における感受性の違いに着目し、音楽が文化や感性と深く結びついていることを多角的に分析する。例えば、西洋音楽における長三和音と短三和音の明暗の対比や、五度圏理論に基づく不協和音の印象などが、必ずしもすべての文化圏で共通の認識を持つわけではない。

また、ドミナントからトニックへの完全終止の効果が聴衆の感動を引き出すメカニズムについて、心理学や神経科学的知見も取り入れつつ考察を深められている。期待が裏切られず適度に叶えられることで生じる快楽の原理を、楽曲分析と対応づける。こうした学際的アプローチによって、楽曲が人間の心理に訴えかける仕組みを多面的に解き明かすことを目指す。

メジャーコードとマイナーコードに焦点を当てそれぞれの機能について説明する。

機能と和声の理論と実践

「コードの違いによる印象の違い」に焦点を当て、音楽理論の中でも機能と和声に注目し、メジャーキーおよびマイナーキーにおける基本的なコードの役割や機能について探究した。六つの基調和音がメジャー3人衆とマイナー3人衆に分かれ、それぞれが異なる印象を生むことが強調され、機能と和声の理論を通じて具体的な役割に分類された。これにより、コード進行の理解が深まり、機能と和声が作曲や演奏においていかに印象を左右するかが示唆された。

六つの基調和音がメジャー3人衆とマイナー3人衆に分けられ、それぞれが異なる印象を生む。この差異が音楽理論において「機能と和声」の概念として取り入れられ、コード進行の理解を深める基盤となる。メジャーキーにおいては、トニック機能(T)であるIが中心に立ち、IVやVが曲を展開させるドミナント機能(D)およびサブドミナント機能(S)を果たす。これにより、同じメジャーコードでも異なる印象が与えられ、聴感的な理解が深まる。

機能と和声はコードの機能に基づいてコード進行を理論的に論じるアプローチであり、トニック機能、ドミナント機能、サブドミナント機能の3つの機能種によってコードを分類する。これにより、コード進行が展開される際の印象や感情の違いが明確になり、理論的な理解が進む。

マイナーキーにおいても、トニック機能、ドミナント機能、サブドミナント機能が存在し、それぞれがメジャーキーとは異なる特徴を持つことが示された。特にIII_mの二面性が強調され、その前後関係や音の配置によって異なる印象を生むことが指摘された。

これらの理論的なアプローチは、作曲や演奏において実践的な価値を持つ。例えば、即興演奏やセッション、アレンジメントの際にリアルタイムでコード譜を作成する際に機能と和声の理解が印象の調整に役立つ。また、機能と和声を基にした理論的な知識は、音楽教育やスタジオレコーディングにおいても印象の演出において鍵となる。

機能と和声は、コードの機能を3つのカテゴリーに分類することで、コード進行における印象や効果を理論的に分析できる枠組みを提供する。具体的には、中心的な機能であるトニック機能、曲の展開を支えるドミナント機能、トニックへの返る地点となるサブドミナント機能の3つに大別される。これによって、例えばメジャーキーのI - IV - V - Iのコード進行では、Iがトニック機能、IVがサブドミナント機能、Vがドミナント機能を担うことで、聴感的な印象が理論的に説明できるようになる。

さらに、マイナーキーにおいても同様の機能分類が適用され、コードの機能的な役割に応じた印象の差異が明らかにされた。特にマイナーではIII_mコードの二面性が注目され、前後のコンテキストによってトニック寄りの印象とドミナント寄りの印象が生まれることが示唆されている。こうした理論的知見は、実際の作曲や編曲、演奏に応用できる実践的な価値を提供している。

また、コードのTDSを変化させていくことは、曲の流れ作りに大きく関わります。TDSの機能が異なるコードへ進んだ場合に起こる曲想を「着地・浮遊・緊張」という表現になぞらえてまとめると、自由派音楽理論では、この内回り(反時計回り)の動きを機能の順行(Prograde Motion)、外回り(時計回り)の動きを機能の逆行(Retrograde Motion)と呼ぶことにしている。

コード	機能名
I	トニック機能(<i>T</i>)
V	ドミナント機能(<i>D</i>)
IV	サブドミナント機能(<i>S</i>)

図 3.4: メジャーコードの構成

コード	機能種名
I VIm	トニック機能(<i>T</i>)
V IIIm	ドミナント機能(<i>D</i>)
IV IIIm	サブドミナント機能(<i>S</i>)

図 3.5: マイナーコードの構成

機能の順行は、 $T \rightarrow S \rightarrow D \rightarrow T$ という向きのサイクルが「緊張と緩和」のセオリーに則った典型的な動きであり、特に *D* から *T* へと進む結びつきは強いと言われている。ジャズ理論は、この *S-D-T* という流れを土台にして理論を発展させている。

機能の逆行、すなわち $T \rightarrow D \rightarrow S \rightarrow T$ のサイクルも心地よいものであり、曲想を豊かにしてくれる存在です。言葉の響きからネガティブなイメージを持つことはなく、順行と逆行は表現において等しく重要な存在である。

結論として、「コードの違いによる印象の違い」を理解することは音楽の表現力を向上させ、実践的な演奏や作曲において有益である。音楽家や学習者は、機能と声の枠組みを通じてコードの印象の微妙な違いを把握し、これを創造的な表現に活かすことが期待される。機能と声は抽象度の高い理論であるため、実際の曲に当てはめていくことでより体感的な理解が深まり、柔軟な応用力が身に付いていくものと考えられる。

提案手法

§ 4.1 現存のコード譜と作成したコードの比較

FFTを用いた音高推定で検出された音階がAメジャーコードの音階構成(A, C#, E)が実際のAメジャーコードと一致するかどうかを判定する。

まず、コード譜をプロットする際の条件を以下に示す。

フレット制約: 和音の形状が演奏しやすいものとなるように、フレットは4フレット以内に制約している。これにより和音が簡潔であり、演奏者にとって取り扱いやすい形状を優先している。

一弦一音制約: 各弦に対しては、一つの音しか出せないという制約をかけている。これにより、和音がクリアで区別可能なサウンドとなる。

これらの条件により、プログラムはメロディーから抽出した音階、Aメジャーコードに相当する和音を出力する。今回はAメジャーコードとBマイナーコードについての比較を行い有効性を示す。

まず、図に今回のシステムでプロットされたAメジャーコードを示す。これをAコードのコードトーンと比較を行う。これはAメジャーコードの第二転回形となっている。第二転回形とは和音の構成音が特定の順序で配置される形状を指す。Aメジャーコードの場合、通常構成ではA-C#-Eですが、第二転回形ではC#-E-Aといった配置になる。

この配置の変化は和音の特徴的な表現方法であり、構成音の順序が異なることで音楽的な響きや表情が変わる。第二転回形により、Aメジャーコードがより独特で鮮明なサウンドを生み出している。このような詳細な比較を通じて、システムが和音の構成を正確に捉え、特に第二転回形を適切に検出していることが確認できる。

次に今回のシステムでプロットされたBマイナーコードを示す。これもBマイナーコードの第二転回形となっている。第二転回形とは和音の構成音が特定の順序で配置される形状を指す。Bマイナーコードは通常、B-D-F#の3つの音から構成されます。しかし、今回のシステムがプロットしたBマイナーコードは、D-F#-Bという音の並び順になっています。これは、Bマイナーコードの第二転回形という変型です。

第二転回形は、コードを構成する3つの音の順序を変えたものです。今回の場合、Bマイナーコードの根音であるBの音が一番上の弦から鳴っているのではなく、真ん中の弦で鳴っています。

この第二転回形では、Bマイナーコード本来の音色や雰囲気が変わって聞こえてきます。根音から始まらないことで、より複雑な響きになり、おもしろみが増します。

以上のように、今回のシステムはBマイナーコードの第二転回形も正確にプロットでき

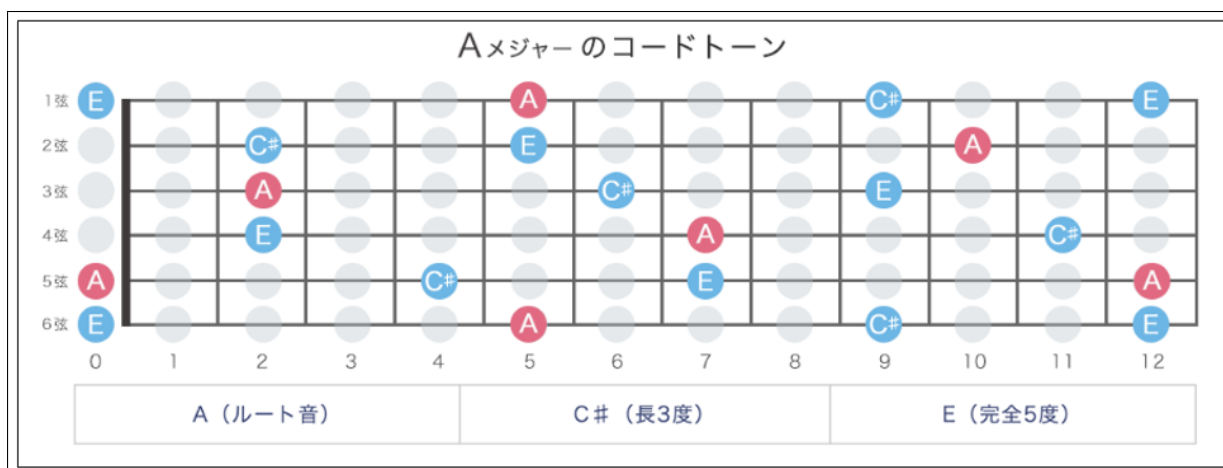


図 4.1: コードトーン

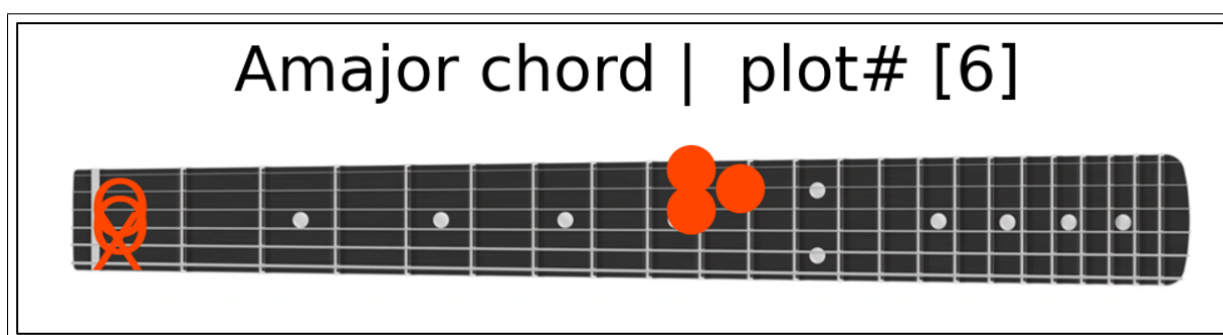


図 4.2: Amajor

ていることが確認できました．コードの微妙なバリエーションまで検出できる性能が実証されたと言えるでしょう．

以上のように，プログラムは複数の条件を基にして和音の検出を行い，特に A メジャーコードと B マイナーコードに焦点を当てながら，その形状や配置に関する詳細な情報を提供しています．これにより，演奏された音楽の解析や理解が深まり，演奏者やリスナーにとって有益な情報が得られることが期待されます．

§ 4.2 タブ情報の保存システム

タブ情報の保存システムは演奏された和音の構造や特性を緻密に記録し，後の詳細かつ総合的な分析を可能にする極めて重要なツールである．このシステムは，単一の和音を構成する極めて多様な要素をほぼ全ての側面にわたって保存することができ，過去に演奏された和音の状態を後から高精度で再現でき，理論的・実践的な考察を行う上で役立つと考えられる．

和音の構成要素として，まず `raw_open_strings` が挙げられる．これは和音の根底となる

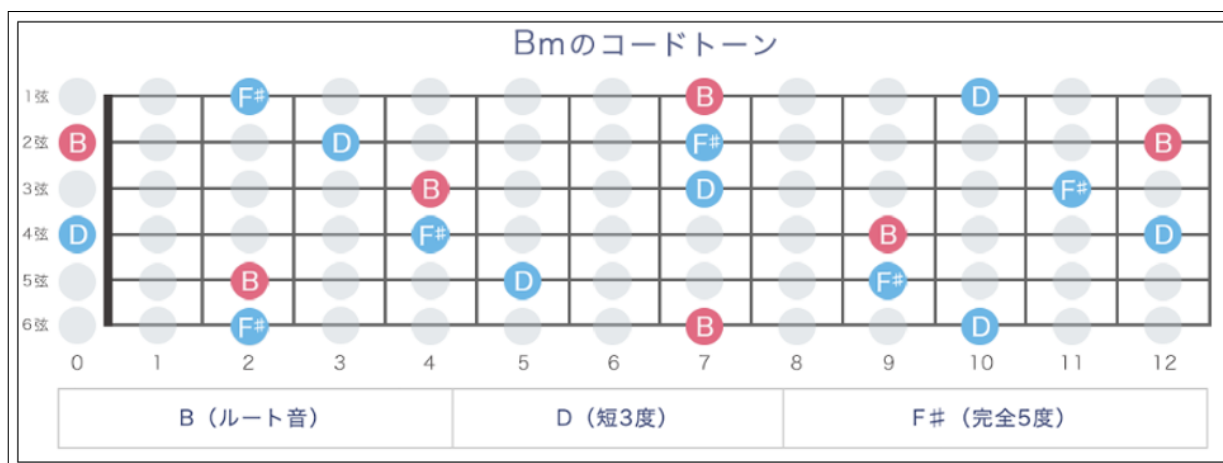


図 4.3: B マイナーコードトーン

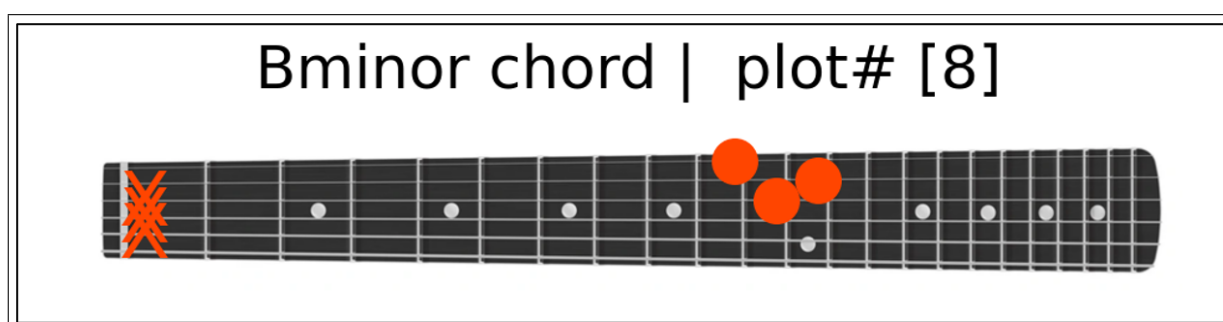


図 4.4: Bminor

基本的な開放弦の位置を示す情報であり、和音全体の出発点となる基本音の供給源たる弦の状態を表している。次に shape があり、これは具体的にはフレット上の指の押さえ方から和音全体の構成音の空間配置に至るまで、和音の実際の形状を詳細に示すフィンガーポジションのデータを指す。この情報は、演奏された和音がどのように成立していたかを立体的に示している。

さらに、演奏時の実際の状態を示す要素として played_strings がある。これは実際に弾かれたり押さえられたりした弦の番号であり、その瞬間に得られた和音の具体的な響き方や実際のフレット上の形状を特定化する情報である。加えて open_strings は演奏を通じて開放された状態を保持し続けた弦の番号を示し、そうした自由弦の基本音が継続的に全体の響きにどのような影響を与えていたかを明示する。さらに muted_strings は抑制され続けた弦のデータであり、演奏の際にどのような特殊な効果や音色の変化を与える処理が施されていたかを示す。

こうした立体的な和音を全方位的に記述するデータを体系的に保存し、後の分析へ提供することで、実際に演奏された和音の詳細かつ全面的な構造や特性を高精度で再現し、理論的実践的に深く理解することを可能にする。要するにこのタブ情報保存システムは、音楽理論研究や作曲制作、楽器教育など、極めて多岐にわたる場面で活用され続けている重

```

raw_open_strings=[1, 3, 4]
shape=[[2, 8), (5, 6), (6, 6)], played_strings=[2, 5, 6]
open_strings=[1, 4]
muted_strings=[3]

raw_open_strings=[1, 3, 6]
shape=[[2, 8), (4, 8), (5, 6)], played_strings=[2, 4, 5]
open_strings=[1]
muted_strings=[3, 6]

raw_open_strings=[2, 3, 6]
shape=[[1, 6), (4, 5), (5, 6)], played_strings=[1, 4, 5]
open_strings=[2]
muted_strings=[3, 6]

```

図 4.5: B メジャーコードの情報

```

raw_open_strings=[4, 5, 6]
shape=[[1, 10), (2, 12), (3, 11)], played_strings=[1, 2, 3]
open_strings=[]
muted_strings=[4, 5, 6]

raw_open_strings=[1, 4, 5]
shape=[[2, 12), (3, 11), (6, 10)], played_strings=[2, 3, 6]
open_strings=[1]
muted_strings=[4, 5]

raw_open_strings=[1, 5, 6]
shape=[[2, 12), (3, 11), (4, 12)], played_strings=[2, 3, 4]
open_strings=[1]
muted_strings=[5, 6]

```

図 4.6: F メジャーコードの情報

要な存在である。特に過去の演奏分析や新曲着想などの文脈で、このシステムが提供する詳細すぎるほどの情報は非常に大きな意味を持っていると考えられる。

§ 4.3 メロディーからコード譜作成システムの開発

本研究の提案システムは、大別すると以下のような4つの工程からなる。

1. ユーザーが演奏を行いたいコードの種類を選択しそれに合ったコードを生成できるようにする。
2. 入力された音声をリアルタイムで取得しFFTを用い、音高判定を行う。
3. 判定された音高に合わせ、ギターの指版に抑える場所を○や×で示したコード譜を作成して画像としてプロットする。
4. 作成されたコード譜にあったタブ譜の情報を保存し共有できるようにする。

これらが本研究のシステムの流れでありユーザーの好みに合わせた演奏形態をとることができる。これらのシステムの流れを説明する。

Step 1: システムの起動とコードの選択

システムを起動してもらい、ユーザー登録とログインをしてもらう。そこでコードの作成、プロットを行う際のコードの種類を選択してもらう。今回はメジャーコードとマイナーコードをはじめとしたセブンスが加わることによりより豊かな音響を持ち、より複雑なハーモニーを作り出す7thスコード、一時的な緊張感を生み出すサスペンド4thとも呼ばれ、サードを持たずに4度が含まれているSus4コードを含めた13種類から選んでもらう。

Step 2: 音声の入力とFFTを用いた音高推定

マイクからのリアルタイムオーディオデータを取得し、その波形と周波数スペクトルを表示する。PyAudioを使用してオーディオデータをストリームから読み取り、Matplotlibを使って波形とFFTをリアルタイムでプロットする。また、音高推定を行い推定された音名を画面に表示する。OpenCVとPILを用いて、日本語フォントを使っ

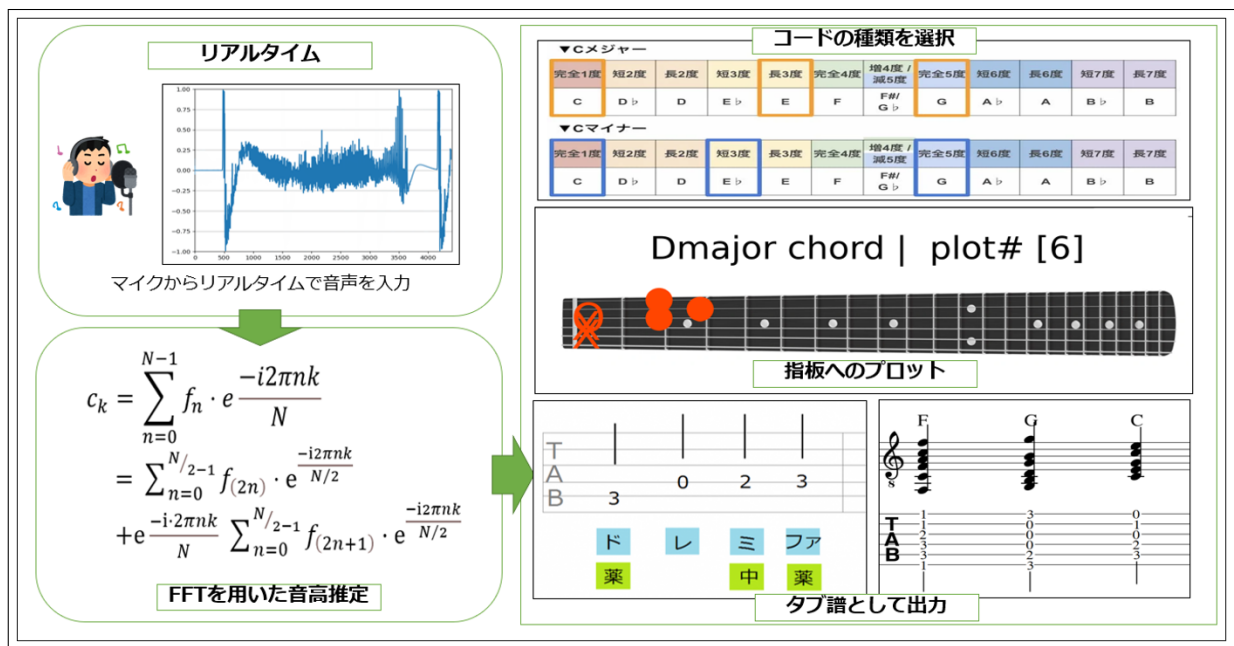


図 4.7: システム全体の流れ

たテキスト描画も行う。プログラムは無限ループで動作し、Ctrl+Cで中断できる。これにより、マイクからの音声を視覚的かつ音楽理論に基づいた形でモニタリングすることができる。

Step 3: 推定された音階に合ったコードの出力

コード譜の出力までの流れについて解説を行う。

コードの音符を取得する

ギターで最も一般的なコードであるトライアドは半音階から来ている。コードを生成するにはルート音を選択し、コードの「品質」に応じて特定の音程を適用して他の2つの音を見つける。たとえば、Amコードの場合Aはルート音mは品質、つまり「マイナー」の標準的な略語である。

コードに属するノートを見つけるには半音スケールから開始する。半音スケールには、それぞれが1つの半音で区切られた12個のノートが含まれている。

A — A♯ — B — C — C♯ — D — D♯ — E — F — F♯ — G — G♯

一部の音符は、A♯/B♭ (シャープ/Bフラット) のように2つの異なる方法で書くことができる。どちらの名前も同じ周波数の音を指すが、異なる文脈で使用されている。

ルート音から始めてマイナーコードの場合は、まずルート音の「短3度」(半音3つ)上の音を取り、Cの音を生成する。3番目の音はルート(7半音)上の「完全5度」でEになる。

つまり、AmコードにはA, C, Eという音が含まれている。

これをコード化するには、まず半音スケール、音程、さまざまな式を表すオブジェクトを追加する必要がある。このクラスは ChromaticScale 特定のルート音のインスタンスを作成し、ルートから始まる半音スケールをキーとしてルートからの距離 (半音単位) を値として持つ辞書を構築する。

フレットボード上で音符を見つける

コード内の構成音がわかったので、次のステップとしてギターのフレットボード上でそれらの具体的な音を見つけ出す必要がある。

ギターのフレット配置は半音階に従っているため、同一弦上で隣接するフレット位置の音は必ず半音の間隔で設定されている。open 弦上を基点として半音ずつ区切られた音がフレット上に並んでおり、最終的に 12 番フレットの位置で元の open 弦と全く同じ音に回帰する。つまりギターの音階は 12 フレット分で 1 オクターブを完結しており、そこからまた新しいオクターブが始まる構成になっている。

この性質を利用して、コードを構成する音が開放弦から何フレット分の位置に存在するかを特定すれば、ギター上での実際の押さえ位置を割り出すことができる。例えば C メジャーのコード音は C, E, G ですが、これらは各弦の 0 フレット、2 フレット、3 フレット上に存在します。したがって C メジャーコードの基本形をこの情報から導き出せる。

コードシェイプの取得

これで、特定のコードを演奏するために考えられるすべての運指位置を計算するための前提条件がすべて揃う。多くの音符で 1 つの弦上に 2 つの異なる位置を設定できるという事実を考慮すると、可能な運指の組み合わせの総数は膨大になる。ただし、いくつかのルールを追加して物理的に再生でき音もおかしくないものだけを取得することが必要である。

音符の間隔は 4 フレットを超えてはいけません。そうしないと、コードを演奏するのが非常に難しくなるか、不可能になってしまいます。一度に発音できる音は 1 つだけであるため、コードのすべての音は別の弦に配置する必要があります。

オープン弦とミュート弦の識別

演奏できないシェイプの一部を除外しましたが、提案されたコードには問題が残っている。ここで、3 つの塗りつぶされた黒い点は指の位置を示し空の円は音をフレットせずに演奏できる「開放弦」を示し、2 つの X マークはミュートする必要がある弦、またはまったく弾かない弦を示している。

4 番目の開放弦は標準チューニングで D 音 (この場合はルート音) を演奏するため、コードの一部である間はフレットを引く必要はない。

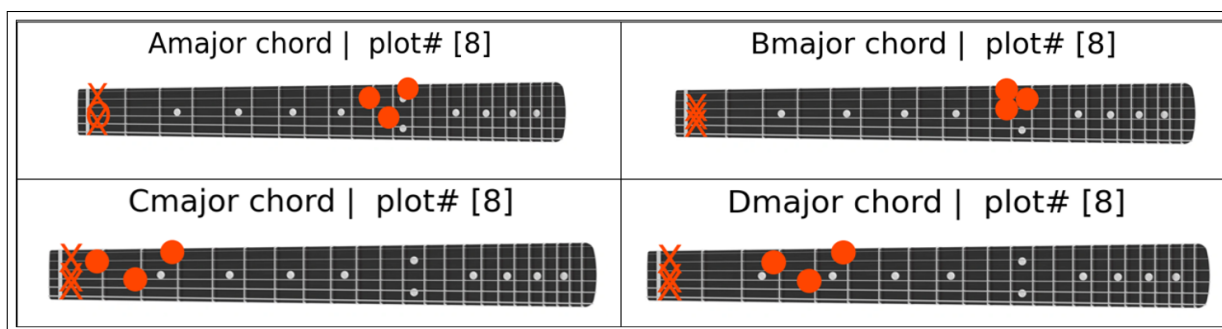


図 4.8: メジャーコード

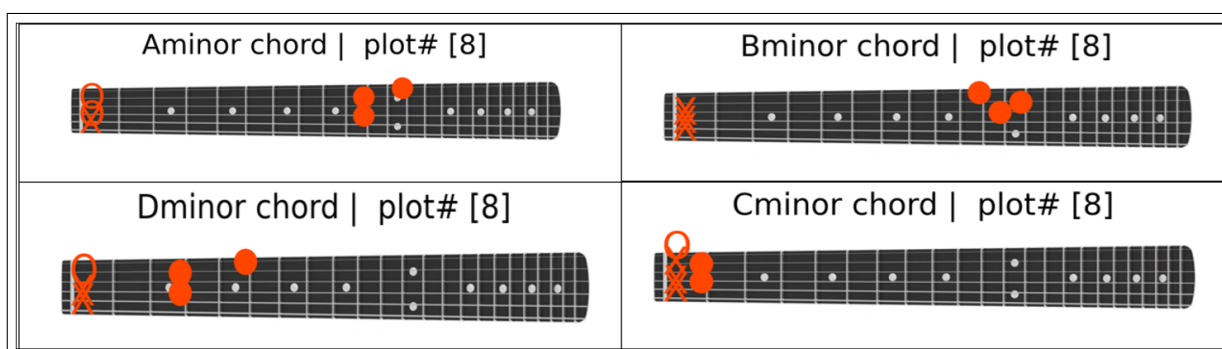


図 4.9: マイナーコード

2本の「下」の弦(「最低音域」という意味では、技術的には上の2つの弦)は、コードを「濁らせ」たり、サウンドを完全に混乱させたりするため、演奏すべきではない。たとえば、5番目の開放弦はA音を演奏する。これは技術的にはコードの一部(D, F#, Aの音で構成されます)だが、ルート音より低い弦にあると「濁ります」耳はルート音をベース(最低音)と結びつける傾向があるためこの音はAになる。専門的には、このコードはD/A(「D over A」または「D with A ベースノート」と呼ばれている。

最も低い弦を開放で演奏した場合E音が生成される。これは音域が根音より低く、コードの一部でもない。ただし、Dメジャー スケールの一部である。したがって、この音が演奏されるとDコードのように聞こえず、代わりにEコードのバリエーションになっている。おそらくE7b9sus4またはE7sus2sus4、正しい名前についての意見は異なっている。

したがって、コードを正しく鳴らすために演奏できる開放弦とミュートすべき弦を識別するメソッドを追加する必要がある。

Step 4: コードに合ったタブ情報を保存

- `raw_open_strings`: 和音を構成する開放弦の位置

- **shape**: 和音の形状を示すフィンガーポジション
- **played_strings**: 演奏された弦の番号
- **open_strings**: 開放された弦の番号
- **muted_strings**: ミュートされた弦の番号

これらの情報を保存することで演奏された和音の構造や特性を後で分析し、理解することができる。音楽理論の研究や楽曲制作、音楽教育などの様々な分野でタブ情報の保存システムは重要な役割を果たしている。

システム上でこれらの4つの Step を処理することによりメロディーに合わせたコード譜の出力を行うことができる。つまりユーザーは音声を入力し、ギターを演奏する処理を求められる。このようなシステムは音楽の即興演奏や耳コピを支援するだけでなく、楽曲の作曲や編曲のプロセスにも役立てることが可能だと考えられる。

数値実験並びに考察

§ 5.1 数値実験の概要

本研究の数値実験としてシステムの有用性の検証、科目推薦の有効性を行う。システムの有用性の検証では実際にシステムを使用してもらい、アンケートに答えてもらう。アンケートの項目は全部で8個あり、その8個には必ず答えてもらう。また、アンケートと同時にコメントを記入できる欄を設けておき、自由にコメントをできるようにする。このアンケートを持って本システムの有用性の検証を行う。アンケート項目は表 5.1 に示す。以上のコメントを5段階のリッカート尺度で評価してもらう。リッカート尺度とは、あるトピックに対して、多段階の選択肢を用いたアンケートを取り、回答者がどの程度同意するか測定する手法である。今回のアンケートでは5段階のうち、1を全く満足していない、2をあまり満足していない、3をどちらでもない、4をやや満足している、5を非常に満足しているといったようにアンケートに答えてもらう。表 5.1 を見てわかる通り、アンケート項目全体を通して、基本的にはシステムの使用感に関する質問を多くしてある。本研究ではアンケート調査を用いてシステムの有効性を示す。調査の対象は同研究室の学部4年、3年生の合計8人に実際に開発したシステムを使用してもらい、アンケートを答えてもらった。実際に使用してもらうにあたり、システムの使用手順について説明を行い、実際に使用してもらう。手順は以下に記してある通り、新規登録から推薦科目の選択、教材のレビュー、ログアウトまでの一連の流れを説明した。

§ 5.2 実験結果と考察

表 5.1: アンケート項目

システムの使い方を理解できたか	このシステムを使用して耳コピの効率が上がると思うか
即興演奏はできたか	自分の考えた音を引くことができたか
コードを正しく抑えることができたか	コードの抑え方を理解できたか
ギターに興味を持つことができた	印象の違いを感じることはできたか

表 5.2: アンケート結果

質問内容	全く満足していない	あまり満足していない	どちらでもない	やや満足している	非常に満足している
システムの使い方を理解できたか	0	0	0	0	0
このシステムを使用して耳コピの効率が上がると思うか	0	0	0	0	0
即興演奏はできたか	0	0	0	0	0
自分の考えた音を引くことができたか	0	0	0	0	0
コードを正しく抑えることができたか	0	0	0	0	0
コードの抑え方を理解できたか	0	0	0	0	0
ギターに興味を持つことができた	0	0	0	0	0
印象の違いを感じることはできたか	0	0	0	0	0

おわりに

本研究では、近年、即興演奏の向上や楽曲再現性の重要性から耳コピ支援システムの研究が進んでいることからアーティストやミュージシャンは耳コピで楽曲を学びその助けとなるのがコード譜を出力するシステムを開発し、その手法を提案した。

まず、システムは入力された音声を取得し、高速フーリエ変換を使用して音高を判定する。次に、ユーザーは演奏したいコードの種類を選択し、それに合ったコードが生成される。判定された音高に基づいて、ギターの指板に押さえる場所を示したコード譜が生成され画像としてプロットされるシステムとした。これにより即興演奏の向上や耳コピの支援を行えるシステムとした。

今後の課題としてはまず音階の判定範囲を現在の1オクターブからさらに拡張し、複数オクターブに対応可能なシステムを構築することが挙げられる。これにより、より幅広い音域のメロディーに対しても正確な音階をリアルタイムで抽出できるようになり、ユーザーが曲調やジャンルにとらわれることなく多種多様な楽曲の耳コピを行えるようになると思われる。また、ギター特有の奏法であるスライド、チョーキング、アルペジオ奏法といった特殊技法の音も識別できるよう信号処理アルゴリズムと学習データの拡張が必要不可欠である。これらの効果音は単純な音階から外れる場合が多く、従来の音階推定システムでは対応が難しい側面がある。こうした特殊技法まで網羅的に処理できるようになれば本システムの有用性と汎用性はいっそう向上することが予想される。

このように、音階識別の範囲と正確性を引き上げるとともに楽器特有の奏法までも柔軟に取り込めるようシステムを進化させていくことが今後の大きな開発課題だと考えられる。これによってユーザビリティの高い汎用的な音楽録音・解析システムへと昇華できるものと期待されている。今後の展望としては音楽理論の導入を行い音楽における教育分野で活用し、耳コピや即興演奏を向上させることができると考えられる。

謝辞

本研究を遂行するにあたり，多大なご指導と終始懇切丁寧なご鞭撻を賜った富山県立大学工学部情報システム工学科情報基盤工学講座の奥原浩之教授，António Oliveira Nzinga René 講師に深甚な謝意を表します．また，システム開発および数値実験にあたり，ご助力いただいた富山県立大学情報システム工学科4年生の〇〇氏に感謝の意を表します．最後になりましたが，多大な協力をしていただいた研究室の同輩諸氏に感謝致します．

2024 年 2 月

山内拓海

参考文献

- [1] 長谷川隆, 西本卓也, 小野順貴, 嵯峨山茂樹, ”楽譜情報からの作曲家らしさ認識のための音楽特微量の提案”, 情報処理学会論文誌 Vol. 53 No. 3 1204–1215 (Mar. 2012).
- [2] 誰でも利用できる音楽視聴支援サービス「Songrium」を一般公開”, 閲覧日 2023-01-025, <https://www.jst.go.jp/pr/announce/20130827>.
- [3] 矢澤一樹, 糸山克寿, 奥乃博, ”ギター演奏音からの難易度調整可能なタブ譜自動生成システム”, 情報処理学会第 76 回全国大会, 5R-5.
- [4] 酒向慎司, ”音響信号によるリアルタイム楽譜追跡と能動的演奏支援システムに関する研究”, 電気通信普及財団 研究調査助成報告書 No. 31 2016.
- [5] 築山裕司, 井上亮文, 市村哲, ”ギタリストの演奏テクニックを視覚化し表現力を向上させる電子楽譜”, 情報処理学会第 69 回全国大会 3Z-7.
- [6] 水野 創太, 一ノ瀬修吾, 白松俊, 北原鉄朗, ”演奏未経験者のためのスマートフォンセンサーを用いた即興合奏支援システムの試作”, The 31st Annual Conference of the Japanese Society for Artificial Intelligence, 2017.
- [7] 平賀悠介, 大石康智, 武田一哉, ”主観評価に基づく楽曲間類似度算出モデル”, 情報処理学会研究報告, Vol. 2009-MUS-81 No. 2, 2009/7/29.
- [8] 張夢堯, 成凱, ”鼻歌検索のための音楽特微量の抽出と評価”, DEIM Forum 2022 C21-4.
- [9] 富山優亮, 泉朋子, ”演奏の楽しさの創出を目的としたギター未経験者向け音再生システムの提案”, 情報処理学会第 81 回全国大会.
- [10] 土橋彩香, 池宮由楽, 糸山克寿, 吉井和佳, ”歌声・調波楽器音・打楽器音分離とユーザ演奏のリアルタイム可視化に基づく音楽演奏練習システム”, 情報処理学会 インタラクション 2016.
- [11] 池田周平, 平賀譲, ”ジャズ即興演奏のための自動伴奏と半自動演奏アプリケーションの検討”, 情報処理学会研究報告, Vol. 2019-MUS-123 No. 16.
- [12] 矢澤一樹, 糸山克寿, 奥乃博, ”ギター演奏者の習熟度に合わせた音響信号からのタブ譜自動生成”, 情報処理学会研究報告, Vol.2013-MUS-100 No.17.
- [13] 株式会社ネオマーケティング, ”リッカート尺度とは？特徴や調査例、メリット・デメリット徹底解説 — ネオマーケティング”, 閲覧日 2022-01-16, <https://column.neo-m.jp/column/marketing-research/-/3617>.
- [14] 後藤真孝, 吉井和佳, 藤原弘将, ”ユーザが誤り訂正により貢献可能な能動的音楽鑑賞サービス”, Proc. IPSJ Interaction, pp. 1363–1372(2012).

- [15] 中野倫靖, 後藤真孝, 平賀譲, ”歌を「歌って/聴いて/描いて」見る歌唱力向上支援インタフェース”, IPSJ Interaction, pp. 195–196 (2007).
- [16] 中村友彦, 水野 優, 鈴木孝輔, 中村栄太, 樋口祐介, 深山 寛, 嵯峨山茂樹, ”音楽演奏の誤りや反復に頑健な音響入力自動伴奏”, 日本音響学会 2012 年秋季研究発表会, pp.931-934, 2012.
- [17] 山本龍一, 酒向慎司, 北村正, ”隠れセミマルコフモデルと線形動的システムを組み合わせた音楽音響信号と楽譜の実時間アライメント手法”, 情報処理学会研究報告, vol.2012-MUS-96(13), pp.16,(2012).
- [18] 山本龍一, 酒向慎司, 北村正, ”セミマルコフ条件付き確率場を用いた音楽音響信号と楽譜のアライメント”, 日本音響学会 2012 年秋季研究発表会, pp.935-936, 2012.
- [19] 後藤真孝, 根本亮, ”遅延を考慮した不特定多数による遠隔セッションシステム”, 情報処理学会論文誌, Vol43,No2. pp299-309,2002.
- [20] 寺田努, 塚本昌彦, 西尾章治郎, ”2つの PDA を用いた携帯型エレキベースの設計と実装”, 情報処理学会論文誌, Vol44,No2,pp266-275,2003.
- [21] 前川督夫, 西本一志, 多田幸生, 間瀬健二, 中津良平, ”ネットワーク型ウェアラブル音楽創奏システムと日常生活空間演出構想の提案”, 日本バーチャルリアリティ学会論文誌 Vol6,No2,pp69-78,2001.