

平成29年度 修士論文

論文題目

コード進行推定に基づき間違いを自動補正する
タッチタイプ可能なVRピアノシステム
Touch-typable VR piano that corrects mistakes
based on chord progression estimation

指導教員

舟橋 健司 准教授

名古屋工業大学大学院 工学研究科 情報工学専攻

平成28年度入学 28414101 番

名前 松末千佳

目次

第1章	はじめに	1
第2章	音楽に関する従来研究	4
2.1	既存のバーチャルピアノ	4
2.1.1	カメラによる運指認識	4
2.1.2	超音波を用いた運指認識	5
2.1.3	3D デプスカメラを用いた運指認識	5
2.1.4	運指認識デバイスの比較	6
2.2	作曲支援に関する研究	7
第3章	タッチタイプでの演奏を実現する手法	9
3.1	提案システムと手法の概要	9
3.2	打鍵位置をもとにした各鍵盤の打鍵確率推定手法	11
3.3	音楽理論をもとにした各鍵盤の打鍵確率推定手法	15
3.3.1	打鍵位置をもとにした曲の調推定手法	15
3.3.2	推定調をもとにした打鍵時のコード推定手法	19
3.3.3	推定調と推定コードを用いた打鍵確率推定	30
3.4	位置打鍵確率と理論打鍵確率を用いた打鍵音確率推定	34
第4章	実験	39
4.1	実験環境	39
4.2	実験を行うにあたっての重み決定	41
4.3	音楽経験者による演奏	46
4.4	音楽未経験者による演奏	48
第5章	むすび	53
	謝辞	55
	参考文献	56
付録A	本論文で用いられる音楽用語	59
A.1	音名, 階名	59
A.2	調, 音程	59
A.3	コード	60
A.4	スケール	61

発表論文リスト

第1章 はじめに

音楽経験者、特にピアノが弾ける人は良いメロディを思いついた時に、そのメロディを確認する目的で無意識に指を動かしてしまう場合がある。身近に楽器があれば実際に演奏を行い音を確認することができるが、身近に楽器がない場合も多い。また、実物のピアノやキーボードは大きいため、普段持ち歩くことは難しい。

これまでに様々なバーチャルピアノに関する研究がされており、バーチャルピアノであれば持ち運び問題の解決も期待できる。例えば、指先の動きをシングルビューカメラで取得し、キーストロークを検出、分析することによって仮想ピアノで音を再現する研究 [1] やプロジェクトで投影した鍵盤上でユーザが演奏を行い、運指をカメラで認識することにより運指の提案を行うシステムも存在している [2]。その他にも医療用の超音波画像を用いて手を認識し、テーブル上での動作を可能とした研究 [3][4] や 3D ジェスチャベースのバーチャルピアノも存在する [5]。その他にも手軽にピアノを演奏できるツールとしてスマートフォンやタブレットで実行できるピアノのアプリケーションが存在している。Apple 社の開発した GarageBand[6] や Bell Standard 社が開発した Touch Piano[7] がその例である。

これらのシステムではいずれもユーザは視覚情報を頼りに鍵盤の位置を把握してバーチャルピアノの演奏を行う。そのためこれらのシステムを用いて演奏する場合には実際の演奏時とは異なり、画面を常に見なければ操作を行うことはできない。音楽経験者の多くは譜面を見ながら楽器の演奏を行ったり、周りの状況を確認しながら演奏することも多いため、手元を確認せずに演奏を行うことが最も自然な演奏方法であると言える。そのため、これらのシステムのように手元を確認しながらでの演奏は、楽器経験者が普段行っている無意識化での指の動きによる自然な演奏とは異なっている。演奏者が手元を見ずに演奏ができる理由の一つは手に力触覚フィードバックがあることであろう。押した鍵盤による鉛直方向のフィードバックのみで

なく、その鍵盤の隣の鍵盤側面に触れることによる水平方向のフィードバックを得ることがある。演奏者は隣の鍵盤との位置関係から自分の指の位置を無意識にかつ正しく判断して次に打鍵すべき位置を決定することで正確に鍵盤を打鍵できるのだろう。

一方で音楽未経験者が作曲を行う場合には音楽理論を勉強するなどの大きな負担がかかる。音楽理論の多くは経験と習慣から組み立てられている。機能と声楽を解いたジャン＝フィリップ・ラモーによる音楽理論書 [8] においても、音楽理論とは伝統的な数と比例に結びついた思考方法を基礎としたものであり、芸術家たちがいわば手探りで見つけ出した和声の規則が、いかに唯一の経験の原理から演繹できるかということだと述べられている。もちろん音楽理論が確立された後に音楽理論に則って作曲された曲も数多く存在するが、音楽理論が確立される以前では作曲者の感性や経験、習慣に基づいて作曲が行われてきた。経験や習慣に基づいて作曲された曲の中には、現存する音楽理論では説明の難しい音列や和声が使われていたり、個人によって解釈の異なる音が使われる場合も存在する。音楽理論の膨大さゆえに、音楽未経験者にとって音楽理論の知識をつけることは大きな負担となり、手軽に音楽理論に基づいた曲を作曲することは困難になることが考えられる。そこで、音楽未経験者の音楽理論を勉強する負担を減らすため、作曲支援に関する研究がいくつか存在する。

例えば旋律とリズムを分割して入力し、メロディを構築する Web ベースのアプリケーションが存在する [9]。このシステムでは音の選択肢をシステムが制限することにより初学者でも不協和音を含まず、違和感のない旋律を生成することが可能である。また、リズムも 1 つのボタンのみを使って入力できるため、初学者でも直感的な入力を行うことができる。この他にもカードに手書きで曲線を描くことにより作曲を行う研究が存在する [10]。描かれた曲線は自動的に音階に沿ってメロディとなり、そのメロディに合わせて伴奏が付加される。これらのシステムでは自然な旋律やハーモニーをつくるため、曲中で使用する音を選択する時に音楽理論を利用して、ユーザの支援を行なっている。

そこで私は音楽理論を利用することにより、鍵盤を見ずに演奏できる、すなわちタッチタイプが可能な VR ピアノを実現できるのではないかと考えた [11]。タッチ

タイプでの演奏を行う場合の問題点は、ユーザの意図しない指のずれによる誤打鍵である。ところで、音楽理論に基づく作曲支援システムにおいて適切な音の並びが生成できるのであれば、繊細な力触覚フィードバックのないVRピアノでもある時点において次に来るべき音の候補を絞り込めるのではないだろうか。音楽理論を用いることにより、誤打鍵した音を音楽的に正しい音に修正することでユーザの思い描くメロディに修正したい。本論文で提案するタッチタイプ可能なVRピアノシステムでは、まずユーザは思い描くメロディを演奏するようにテーブル上で手を動かす。システムはデプスカメラやモーションセンサによりテーブル上の運指を取得する。このとき、ユーザの意図しない指の位置ずれ、すなわち誤打鍵があるかもしれない。しかし、これまでのVRピアノシステムは、これが誤打鍵であるかどうか判断することはなく、打鍵された音を忠実に提示する。本研究で提案する手法では、音楽理論によりこのずれを補正することで、本システムは思い描かれたであろう本来のメロディを提示する。また提案システムでは、ユーザの打鍵位置をどのくらい信頼するか、あるいは音楽理論をどれだけ重要視するか、重み付けを調整できるようにする。これは、音楽は時には、これまでのオーソドックスな音楽理論から離れた新しいメロディが望まれることもあり、この時は取得された打鍵位置を通常よりも強く信頼するためである。一方で、音楽理論の信頼度をより高くした場合には、ユーザがほぼランダムな運指を行った場合でも、その全体的な指の動きを緩やかに考慮した、そして音楽理論に基づいた適切なメロディを生成することができる。すなわちタッチタイプ可能なVRピアノは、未経験者でも弾けるピアノとしても、また作曲支援アプリケーションとしても使用することが可能である。

第2章 音楽に関する従来研究

本章では前章にていくつか紹介したバーチャルピアノに関する研究と作曲支援を行うための研究について紹介する。

2.1 既存のバーチャルピアノ

バーチャルピアノを実現するにはユーザの指の動きを取得することが必要となる。本項ではバーチャルピアノを実現するための運指認識の手法について3種類紹介し、その手法を取り入れた研究についてそれぞれ紹介する。

2.1.1 カメラによる運指認識

運指の取得を行う最も手軽な手法として、一般的なカメラやウェブカメラを用いる方法がある。

Chia-Hung らによるシステム [1] では、一般的に家庭でも使われるようなウェブカメラを用いて指の認識を行なっている。カメラによって撮影された2次元情報に対して画像処理を行うことによって手指の範囲を特定する。従来の手法では手指の位置の検出を行う場合、テンプレートマッチングを行うことが多いが、Chia-Hung らによるシステムでは、まず手指の範囲から輪郭を抽出しグラフ化を行う。そのグラフの増減の傾向を読み取ることで指の先端の位置を取得し、位置が急激に下方に動き、止まった場合に打鍵したと判断する手法をとっている。

竹川らによるシステム [2] は各指の爪に貼り付けたカラーマーカーをウェブカメラで読み取り、画像処理を行うことで運指の取得を行なう。システムは鍵盤投影用プロジェクタ、運指認識用のウェブカメラ、処理用PC、投影用のプラスチック板、MIDI鍵盤によって構成されている。ユーザが既存の曲の演奏をプロジェクタで投影された鍵盤上で行うと、あらかじめ設定された演奏曲の正解運指とユーザの運指

を比較した結果がプロジェクタで重畳される。実験では一般的な光るピアノ鍵盤で提示される運指とシステム提示される運指を用いた演奏の比較を行い、システムを使用したほうが早く演奏を習熟できるという結果を得ている。

Adam らによるシステム [12] ではウェブカメラを用いて実際の鍵盤の領域を認識し、手によって遮蔽された鍵盤を打鍵音とする手法を取っている。キーボード識別にはバイナリ閾値処理、Sobel 演算子、ハフ変換、そして提案アルゴリズムを使用し、個々の鍵を検出するためにまずピアノキーボードに似た領域を発見する。この手法をとることにより、カメラのぶれに耐性のあるシステムの構築が可能である。また、実際の打鍵の認識は今後の課題となっている。

2.1.2 超音波を用いた運指認識

Mikel ら [3] や Claudio ら [4] によるシステムでは医療用の超音波画像を用いて運指の認識を行なっている。システムでは、指の力や手の位置を細かく制御することができ、センサに触れることなく演奏することが可能である。打鍵認識には高速衝突検出を使用する。衝突検出では 1ms 以内に衝突力を計算するため、ハプティックレンダリングに適した Voxmap-Pointshell アルゴリズム [13] を用いている。運動方程式を解くために物理エンジン Bullet [14] を利用し、衝突検出フレームワークの組み合わせにより、複数の複雑な仮想オブジェクトとの相互作用を可能にしている。

2.1.3 3D デプスカメラを用いた運指認識

Enkhtogtokh らは 3D ジェスチャベースのバーチャルピアノの開発を行なっている [5]。Leap Motion を用いて運指の取得を行なっており、y 軸方向の指の速度差が閾値以上である場合、打鍵したと判定する。また、誤打鍵を回避するため、手のひらの移動速度が閾値以下の場合には手の位置が固定されていると判定する。Inwook らによるシステム [15] においても手指の取得を Leap Motion を用いて行っている。3D 空間内に描画された鍵盤に、Leap Motion で取得した指先が触れた場合に打鍵したと判断する。鍵盤と指が描画された 3D 空間はヘッドマウンドディスプレイによって表示される。実際の演奏を行うテーブル上にはハプティックディスプレイが設置

されており、ハプティックディスプレイからは40kHzの超音波が出力されることにより、ユーザーに力触覚フィードバックを与えることが可能である。しかし、力触覚フィードバックは垂直方向のみであり、水平方向のフィードバックはない。すなわちこのシステムでは鍵盤の中央を打鍵したのか、端を打鍵したのかが触覚情報から判断できるわけではなく、すなわち鍵盤全体に対する詳細な自分の手の位置が認識できないため、誤打鍵を防ぐことはできない。

2.1.4 運指認識デバイスの比較

Yuriらによる論文で、上記の運指認識手法の比較実験がされている。Yuriらによるシステム [16] ではタブレットの背面に設置されたウェブカメラにより演奏中の画像を取得し、タブレット本体で指の鍵盤を重畳表示することによってVRピアノを実現している。しかし、現在のシステムでは画面上での遅延が存在している。ユーザーによるシステムの遅延の認識は、触覚のフィードバック、聴覚のフィードバック、視覚のフィードバックを行うことにより緩和することができると考えられる。論文内ではより良いピアノ鍵盤インタフェースの実現を目的とし、マルチモーダルフィードバックの知覚を定量化するための2種類の手法の比較実験を行なっている。赤外線カメラが内蔵されたデバイスによる手指認識手法によるシステムと、超音波マトリクスによる手指認識手法によるシステムを構築し、検出位置、ハードウェア要件、fps、誤検出、検知漏れ、待ち時間、触覚フィードバック、被験者の対応、Androidとの統合手法の8項目において比較を行なった。超音波マトリクスによるシステムはHoshiら [17] やCarterら [18] の研究をもとに作成された。それぞれの結果は以下の通りである(表2.1)。

実験後に現段階のLeap Motionは垂直方向の動きの検出には不向きであることや、俊敏な動きを行なった後に指を認識できない場合があることが示されている。また、Leap Motionは赤外線によって手指の検出を行う為、太陽光により検出を阻害される場合が多いことや、手を識別する為に多くのCPU時間を必要とすることが問題点として述べられている。一方で超音波マトリクスによるシステムはより正確な運指の取得を行うことができたものの、設置の段階で多くの時間を有する為、ポータブルなユーザインタフェースではないものの、商業的な実装は可能であると述べら

表 2.1: Leap Motion と超音波マトリクス検出の比較

項目	Leap Motion	超音波マトリクス
検出位置	連続的な検出	1 鍵盤につき 1 つのセンサ
ハードウェア要件	ラップトップ型の Android	Arduino 経由の Android
fps	30 300	200
誤検出	非常に多い	少ない
検知漏れ	少ない	無し
待ち時間	許容範囲内	優れている
触覚フィードバック	離散空間 (テーブルタップ)	連続 (メッシュ上でのタップ)
被験者の反応	悪い	良い
統合手法	未実装	Arduino ADK

れている。

2.2 作曲支援に関する研究

作曲は音楽理論が必要であることから、音楽初心者にとって難易度が高い。そのため、音楽初心者による作曲活動を支援するためのシステムが多く発表されてきた。本項ではこれらの作曲支援を行うための研究について紹介する。また、各音楽用語は末尾の付録 A にまとめた。川島らによるシステム [9] では旋律とリズムを別々に入力することによって直感的に作曲を行える環境を実現している。まずユーザは「明るい：長調」「暗い：短長」や「民俗調」「和風」などの雰囲気を選択する。システムは選択された雰囲気に合ったスケールの中から 4～5 音を提示する。音を提示する際には不協和音にならないように音を選択するため、ユーザはどの音を選択しても不協和音のない音を用いた作曲を行うことができる。Yoshida らによるシステム [19] ではまずユーザが複数の悲しい、楽しいなどの印象語を選択した後、メロディアウトラインを描画することによって旋律を生成する。絵を描くように点を打ち、ラインでつなぐことによってメロディを生成することができる。音楽理論として不協和音を回避するための処理や、コード進行を考慮した処理が行なわれている。このシステムでは不協和音を回避するために、コードトーン以外の音を不協和音として扱う。コードトーン以外の音を選択された場合、選択した音から一番近いコードトーンに音をシフトする。コード進行の部分では石坂によるコード進行パターンの例 [20] に

従い、8小節のコード進行を決定する。また、メロディは全て調号のないCメジャーかAマイナーのどちらかに振り分けられる。Kitaharaらによるシステム[21]では、ユーザが描画したメロディアウトラインに対し、機械学習データによって修正を行うことで作曲の支援を行っている。ユーザの描画したメロディと音楽理論に最も則したメロディとの妥協点から最終的なメロディを決定する。音楽理論に即したメロディは論文内では手動で設定されている。Iwataらのシステムでは[10]アナログのカードにメロディアウトラインを描画し、それをウェブカメラで読み取り、画像処理を行うことによってメロディを生成する。メロディ生成にはLazy Composer[22]を用いており、ユーザが思い浮かべたメロディの音程や長さを曲線で入力することで簡単に作曲ができる。Lazy Composerでは川島らによるシステム[9]と同様に、描いたメロディラインを「長調」「短長」などの調性の設定を行うことにより同一のメロディでも異なるメロディを生成することができる。

第3章 タッチタイプでの演奏を実現する手法

本章では打鍵音の推定手法について述べる。また音楽理論としての、打鍵音推移と現在の打鍵位置に基づく曲の調と当該小節に適用するコードの推定手法について説明する。

3.1 提案システムと手法の概要

テーブル上でVRピアノの操作をする場合、指の動きを取得することが必要となる。タブレット上に表示された鍵盤を打鍵することは簡単であるが、打鍵可能な範囲がタブレットのサイズによって制限されるため、演奏の自由度は低くなってしまふ。タブレット上で画面外の鍵盤を打鍵する場合には画面上で鍵盤をスライドさせる等の別の動作が必要となってくるため、ある程度の自由な演奏を行う場合においての有用性は高くない。本提案手法を適用するシステムではLeap Motion[25]を用いて指先の位置と手の関節を取得する。なお、通常のカメラやタッチパネルを用いた手指の位置の取得方法においても、この提案手法は適応可能である。また本システムでは鍵盤の画像をテーブル上には表示しない。音楽経験者は実際のピアノであれば指の位置を確認することなく演奏を行うことができる。しかし、鍵盤からの物理的なフィードバックがない場合、ユーザは鍵盤の中央を打鍵したのか、端を打鍵したのかを触覚情報から判断できない。すなわち、2章で紹介したこれまでのVRピアノでは、鍵盤全体に対する自分の指の位置を認識して、無意識に指の位置を補正し、次の鍵盤を正しく打鍵することは困難である。VRピアノにおいて実際に取得した運指データのみで打鍵音の決定をした場合、その音の、ユーザが頭の中でイメージした音に対する信頼性は高くない。時には指が打鍵しようとした鍵盤から外れてしまう場合もある。そこで、打鍵された鍵盤位置のずれを音楽理論により修正して最終的な打鍵音を決定する手法を提案する。ここで打鍵音とは最終的に実際に

奏でる音のことであり，また以下で用いる音楽用語は付録 A を参照されたい。

具体的にはまず取得された打鍵位置から近傍の各鍵盤の確率 (打鍵位置確率) を推定する。打鍵位置とは，鍵盤座標系における打鍵時の指先位置とする。次にその確率に基づいて，演奏している曲の調を推定する。曲の調の推定は，川島 [9] や池田 [22] らの手法で用いられているスケールの構成音をもとにした推定手法や，サンプルデータをもとにした統計的な推定手法に基づいて行う。次に推定した曲の調 (推定調) を用いて打鍵時のコードを推定する。コードの推定は Yoshida ら [19] の研究で用いられているコードトーンに基づいた推定手法や，調の推定と同様のサンプルデータをもとにした統計的な推定手法に基づいて行う。最後に，推定したコード (推定コード) に基づく各鍵盤の確率 (理論打鍵確率) と，位置打鍵確率から打鍵音を決定する。演奏の進行に従い打鍵音の推移を蓄積しておき，調が改めて再推定された場合，すなわちそれまでの推定結果が修正された場合にはこれまでこれまでの打鍵音推移データを用いてコードも再推定することにより，ユーザの思い描いたメロディにより近いメロディの生成を行う。

3.2 打鍵位置をもとにした各鍵盤の打鍵確率推定手法

ユーザが打鍵したい鍵盤として選択される鍵盤は、打鍵位置に近い鍵盤でなければならない。また、打鍵位置と打鍵音の関係が信頼できるものでなければならない。そこで、 n 回目 ($n > 0$) の打鍵が行われた時刻 t_n における打鍵位置 (x_n, y_n) に対して、底辺の長さが $2|x_{n-1} - x_n|$ 、面積が1である二等辺三角形を考える (図3.1)。この三角形を分割し、各鍵盤あたりの面積を求めることにより、 n 回目の打鍵時に鍵盤 S が打鍵された確率 $P_{pos_{S,n}}$ (S は3オクターブ分の C, Db, D, Eb, E, F, Gb, Ab, A, Bb, B) の推定を行う (図3.2)。

白鍵と黒鍵の打鍵の判定は、鍵盤上の y 座標を用いて行う。黒鍵と白鍵の境界の y 座標を Y とする (図3.1)。打鍵時に $y_n > Y$ の場合、黒鍵及び白鍵の奥の部分被打鍵したと判定する。また、 $y_n \leq Y$ の場合には白鍵を打鍵したと判定する。前述の二等辺三角形は、以下の3種類の場合に分けて境界線 (ボーダー) により分割する (図3.2)。まず $y_{n-1} \leq Y \cap y_n \leq Y$ の場合は、図3.2 (1) のように白鍵の幅ごとにボーダーを設定する。 $y_{n-1} \leq Y \cap y_n > Y$ の場合、手指が白鍵から白鍵に移動していることからユーザには黒鍵を打鍵する意思がないと考えられるため、黒鍵の幅を0とし、白鍵の幅と同様の幅で設定したボーダーで $P_{pos_{S,n}}$ を決定する。続いて $y_{n-1} > Y \cap y_n \leq Y$ の場合、図3.2 (2) のようにボーダーを設定する。 $y_{n-1} > Y \cap y_n > Y$ の場合、黒鍵から白鍵に手指が移動していることから、ユーザが演奏している曲の調に黒鍵が含まれていることが推定できる。そのため、まず黒鍵と白鍵の境界ごとにボーダーを考える。そして現在打鍵した位置が $y_n \leq Y$ であることから、ユーザには白鍵を打鍵する強い意思があると推定できる。そのため黒鍵の幅を少し狭めた位置にボーダーを設定する。最後に、 y_{n-1} に関わらず $y_n > Y$ の場合、図3.2 (3) のようにボーダーを設定する。 $y_{n-1} > Y \cap y_n \leq Y$ の場合と同様にまず、黒鍵、白鍵それぞれの鍵盤ごとにボーダーを設定する。 $y_n > Y$ の場合には、ユーザには黒鍵を打鍵する強い意志があるものと推定できるため、黒鍵の幅を少し広げた位置にボーダーを設定し、 $P_{pos_{S,n}}$ の決定を行う。ここで n 回目の打鍵に対する最も高い確率 $P_{pos_{S,n}}$ を P_{val_n} とし、またそのときの音 S を位置に基づく打鍵音候補 (以下、仮打鍵音) S_{pos_n} とする。

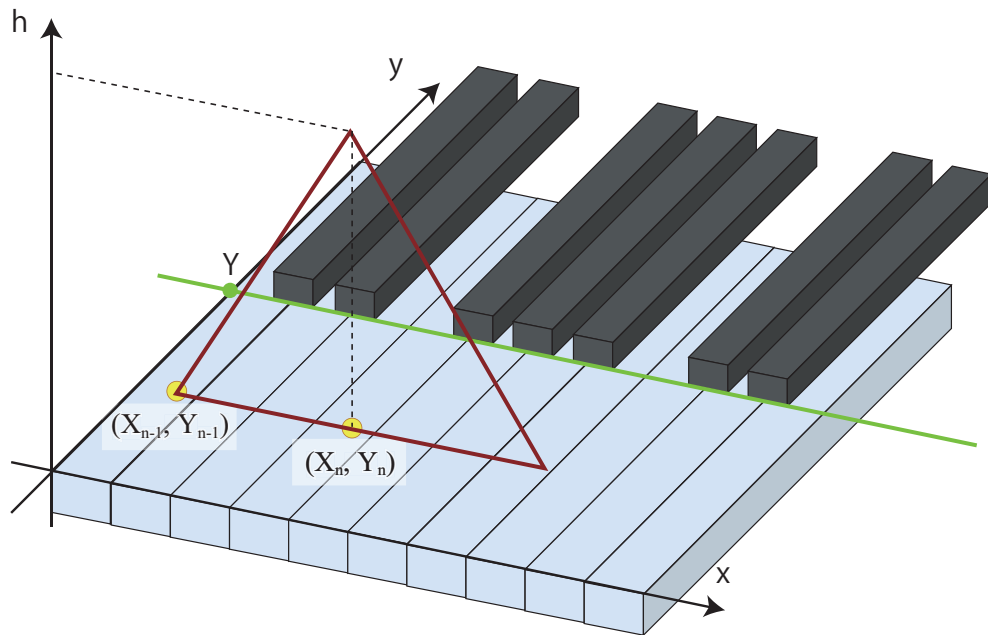


図 3.1: 鍵盤の配置

例えば、 $n-1$ 回目の打鍵が D 、 n 回目の打鍵が G 、 $y_{n-1} \leq Y \cap y_n \leq Y$ であるとき (図 3.2 (1))、確率が 0 ではない鍵盤は D, E, F, G, A, B, C の 7 音であり、それぞれに対して $P_{pos_{D,n}}, P_{pos_{E,n}}, P_{pos_{F,n}}, P_{pos_{G,n}}, P_{pos_{A,n}}, P_{pos_{B,n}}, P_{pos_{C,n}}$ を考える。三角形全体を前述のボーダーにより分割してそれぞれの面積として各鍵盤の打鍵確率を決定する。最も確率の高い G の音を仮打鍵音 S_{pos_n} とする。また、 x_n と x_{n-1} の距離 d が近い場合には遠い場合と比較して各鍵盤の打鍵確率は高くなる (図 3.3)。すなわち指や手全体の移動が少ない場合には誤打鍵は起こりにくいため、後述の音楽理論による判断よりも実際に取得した打鍵位置を信頼して最終的な打鍵音 S_n が決定できる。

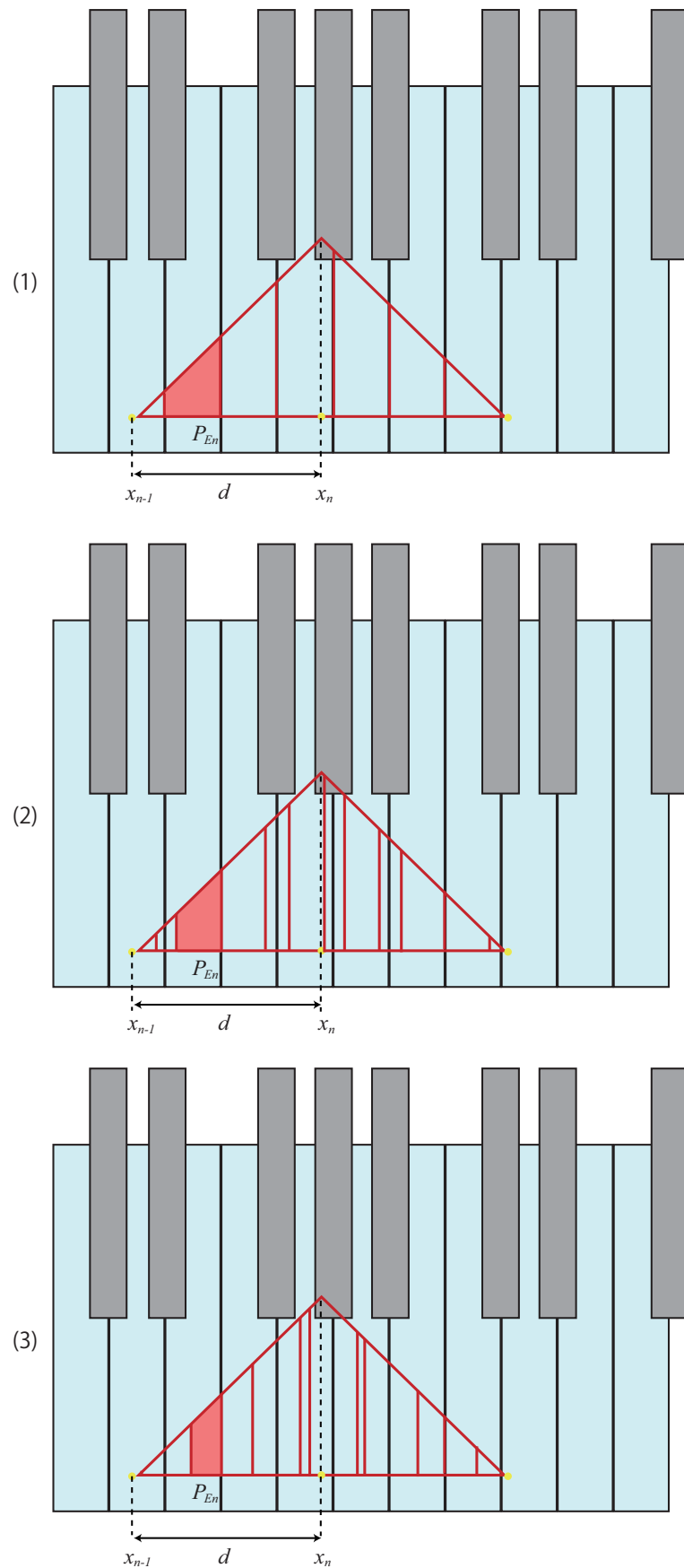


図 3.2: 打鍵位置をもとにした打鍵音推定手法 (黒鍵と白鍵)

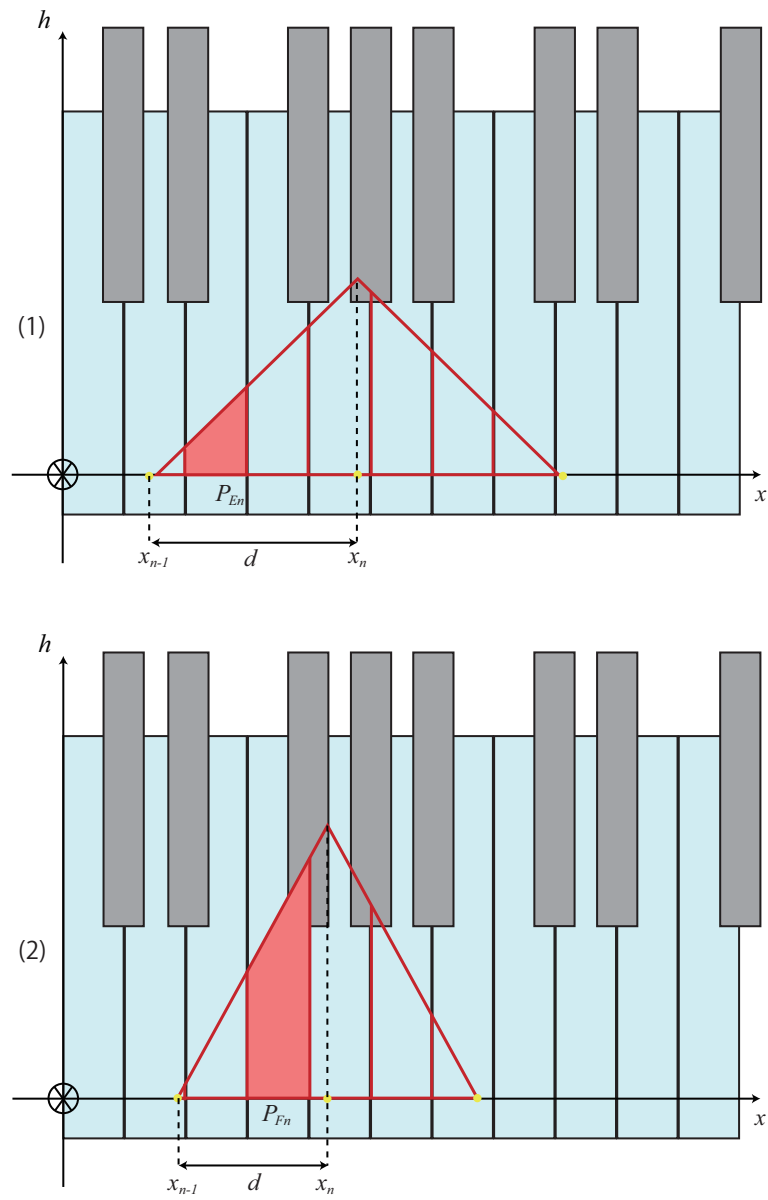


図 3.3: 打鍵位置をもとにした打鍵音推定手法 (指移動距離の違い)

3.3 音楽理論をもとにした各鍵盤の打鍵確率推定手法

この節では音楽理論をもとに各鍵盤の打鍵確率を推定する手法について説明する。まず最初に節 3.2 に述べた n 回目の打鍵音に対する仮打鍵音 S_{pos_n} を利用して、演奏している曲の調 key_n を推定する。次に推定した調 key_n をもとに時刻 t_n のときのコード $code_n$ の推定を行う。

3.3.1 打鍵位置をもとにした曲の調推定手法

演奏曲の調 key_n は、時刻 t_n においての調 k であると判断する確率 (調確率) $Pkey_{k,n}$ のうち、最も確率が高い調 k とする (k は C, Db, D, Eb, E, F, G, bG, Ab, A, Bb, B の 12 種類)。時刻 t_n においては、1 回目から $n-1$ 回目までの $n-1$ 個の打鍵音 S_i (S_1 から S_{i-1}) は既に奏でられており、 key_{i-1} も推定済みであると考え。時刻 t_n において調 k である調確率 $Pkey_{k,n}$ は、1 回目から $n-1$ 回目までの $n-1$ 個の打鍵音 S_i の推移と n 回目の仮打鍵音 S_{pos_n} が、ある調 k のダイアトニックスケール [A.4] の構成音である回数 (個数) が考慮されるように決定する。なお、演奏の途中で曲の調が変化することはないものとする。しかし、打鍵を繰り返すうちに、打鍵音推移に仮打鍵音 S_{pos_n} を加えることで、打鍵音推移が調 key_{n-1} のダイアトニックスケールの構成音である回数よりも他の調のダイアトニックスケールの構成音である回数の方が上回る場合がある。すなわち、過去の調推定が正しくなかったと判断される場合には、新たな推定調 key_n は key_{n-1} と異なるものに更新する。ところで、時刻 t_1 において 1 回目の打鍵を行った時点では、既に奏でられた打鍵音推移はなく、仮打鍵音 S_{pos_1} の 1 音だけでユーザのイメージする曲の調推定を行うことは非常に困難である。そこで情報不足を補うために、既存の楽曲を参考に 1 回目の打鍵における調推定の正解率を高めることを試みる。本研究では「ジャズ・スタンダード・バイブル～セッションに役立つ不朽の 227 曲～」[24] (ジャズバイブル) の収録曲の中から、よく演奏される 100 曲を選出し、それをサンプルデータとして扱う。曲はメジャーキー、マイナーキー、ブルースの 3 種類に分類し、現時点ではその中からメジャーキーの 71 曲をサンプルデータとして利用する。マイナーキー、ブルースの曲の考慮は今後の課題とする。

調の推定を行うためにサンプルデータから2種類の集計を行う。1つ目は全てのサンプルデータ曲を調ごとに分類し、その割合をまとめたデータセットである(表3.1)。この分布から、実際にユーザが演奏を行うときに、実際に存在する曲の中で割合の少ない調の曲を演奏する可能性は少ないと推測し、調を推定する。このデータセットより決定する調 k ごとの確率を $Pkey-data_k$ とする。2つ目は曲の第一音目の音 s ごとに曲の調を分類し、その割合を調べたデータセットである(表3.2)。例えばサンプルデータ中でFの音で始まる曲の場合、調がEbメジャーである曲が全体の16.7%、Fメジャーである曲が全体の33.3%、Bbメジャーである曲が全体の50%、その他の調は0%である。このデータセットより決定する調 k ごとの確率を $Pkey-first_{k,s}$ とする。また、1音目に対しては過去の打鍵音推移は存在しないが、ユーザが当該システムを既に利用している場合などに、そのユーザの好みや嗜好を反映することも将来的に検討したい。そこで、便宜的に時刻 t_0 における調確率 $Pkey_{k,0}$ を考える。なお特に情報のない場合には、いずれの調も平等に選択される可能性があるため、すべて同じ値 ($100/12=8.3\%$) とする。

時刻 t_1 における各調 k の調確率 $Pkey_{k,n}$ は以下の式3.1によって求める。

$$\begin{aligned} Pkey_{k,1} &= Pkey'_{k,1} \times Wkey + Pkey-data_k \times Wdata + Pkey-first_{k,s} \times Wfirst \\ Pkey_{k,n} &= Pkey'_{k,n} (n \geq 2) \end{aligned} \quad (3.1)$$

ここで $Wkey$, $Wdata$, $Wfirst$ は加重平均を求める際の重みであり、 $Wkey+Wdata+Wfirst=1$ である。また、 $Pkey'_{k,n}$ は以下の通りである。

$$Pkey'_{k,n} = Pkey''_{k,n} + \left(1 - \sum_{i=C}^B Pkey''_{k,i}\right) \times \frac{count_{k,n}}{7n} \quad (3.2)$$

$$Pkey''_{k,n} = \begin{cases} Pkey_{k,n-1} & (Spos_n \text{ が調 } k \text{ の構成音の場合}) \\ Pkey_{k,n-1} \times 0.5 & (Spos_n \text{ が調 } k \text{ の構成音でない場合}) \end{cases}$$

ここで $count_{k,n}$ は、時刻 t_{n-1} までの打鍵音推移と時刻 t_n での仮打鍵音 $Spos_n$ が、調 k のダイアトニックスケールの構成音に含まれている回数値である。仮打鍵音がダイアトニックスケールの構成音ではない場合において調確率 $Pkey_{k,n}$ に0.5を乗じるのは、仮打鍵音が調 k のダイアトニックスケールではない場合には推定調が k で

表 3.1: 曲の調の分布 (%)

調	C	D \flat	D	E \flat	E	F	G \flat	G	A \flat	A	B \flat	B
割合	21.13	0	0	15.49	0	28.17	0	8.45	7.04	0	19.72	0

ある可能性を低く考えるためであり、また意図的に弾いた装飾音符やユーザの単純な運指ミスも考慮して可能性を完全に否定しないためである。この操作における減分が式 3.2 の 1 行目の括弧内である。ところで全ての音は 12 種類の調のうち、7 つの調のダイアトニックスケールに構成音として含まれる。すなわち、時刻 t_n における各調 k に対する $count_{k,n}$ の総和は $7n$ である。そこで、この減分を $count_{k,n}$ の値に従い、各調に傾斜配分する。これにより得られた $Pkey_{k,n}$ のうち、最も確率が高いときの値を $Kval_n$ とし、調 k を時刻 t_n のときの調 key_n と決定する。

以下に 1 回目の打鍵が A \flat 、2 回目の打鍵が D \flat であったときの例を示す。まず、時刻 t_0 のときの調確率 $Pkey_{k,0}$ を表 3.3 (1) に示す。A \flat を打鍵した場合、A \flat がダイアトニックスケールに含まれる調は、D \flat 、E \flat 、E、G \flat 、A \flat 、A、B のメジャーの 7 種類である。ダイアトニックスケールの構成音に含まれない調 k において式 3.2 を用い、 $Pkey'_{k,1}$ を決定する (表 3.3 (2))。なお $count_{k,n}$ は表 3.3 (3) となる。式 3.2 により調確率 $Pkey'_{k,1}$ を決定する (表 3.3 (4))。また、 $Pkey-data_k$ はサンプルデータより表 3.3 (5) との値を用い、第一音目が A \flat であることから $Pkey-first_{k,A\flat}$ は表 3.3 (6) の値を用いる。 $Pkey'_{k,1}$ と $Pkey-data_k$ と $Pkey-first_{k,A\flat}$ と $Pkey-data_k$ を 5:3:2 で加重平均をとった場合、調確率 $Pkey_{k,1}$ は表 3.3 (7) となる。

次に、続けて 2 回目に D \flat を打鍵したときの例を表 3.4 に示す。1 回目打鍵時の調確率 $Pkey_{k,1}$ を D \flat を表 3.4 (1) に示す。D \flat を打鍵した場合、D \flat がダイアトニックスケールに含まれる調は D \flat 、D、E、G \flat 、A \flat 、A、B のメジャーの 7 種類である。よって、 $Pkey'_{k,2}$ は表 3.4 (2) となり、 $count_{k,2}$ は表 3.4 (3) となる。減分を傾斜配分した後の調確率 $Pkey_{k,2}$ は表 3.4 (4) となる。

表 3.2: 曲の第一音に対する調の分布 (%)

第一音	C	D \flat	D	E \flat	E	F	G \flat	G	A \flat	A	B \flat	B
C	28.6	0.0	0.0	7.1	0.0	50.0	0.0	0.0	14.3	0.0	0.0	0.0
D \flat	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
D	7.1	0.0	0.0	0.0	0.0	7.1	0.0	42.9	0.0	0.0	42.9	0.0
E \flat	0.0	0.0	0.0	66.7	0.0	0.0	0.0	0.0	33.3	0.0	0.0	0.0
E	100	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
F	0.0	0.0	0.0	16.7	0.0	33.3	0.0	0.0	0.0	0.0	50.0	0.0
G \flat	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
G	55.6	0.0	0.0	11.1	0.0	11.1	0.0	0.0	11.1	0.0	11.1	0.0
A \flat	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100	0.0	0.0	0.0
A	9.1	0.0	0.0	0.0	0.0	81.8	0.0	0.0	0.0	0.0	9.1	0.0
B \flat	0.0	0.0	0.0	66.7	0.0	0.0	0.0	0.0	0.0	0.0	33.3	0.0
B	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

表 3.3: 1回目に A \flat を打鍵したときの $Pkey_{k,1}$ (%)

調	C	D \flat	D	E \flat	E	F	G \flat	G	A \flat	A	B \flat	B
$Pkey_{k,0}(1)$	8.3	8.3	8.3	8.3	8.3	8.3	8.3	8.3	8.3	8.3	8.3	8.3
$Pkey_{k,1}^{\triangleright}(2)$	4.2	8.3	4.2	8.3	8.3	4.2	8.3	4.2	8.3	8.3	4.2	8.3
$count_{k,1}(3)$	0	1	0	1	1	0	1	0	1	1	0	1
$Pkey_{k,1}'(4)$	4.2	11.3	4.2	11.3	11.3	4.2	11.3	4.2	11.3	11.3	4.2	11.3
$Pkey-data_k(5)$	21.1	0	0	15.5	0	28.2	0	8.5	7.0	0	19.7	0
$Pkey-first_{k,A\flat}(6)$	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100	0.0	0.0	0.0
$Pkey_{k,1}(7)$	5.5	3.4	1.3	6.5	3.4	6.9	3.4	2.9	54.8	3.4	5.2	3.4

表 3.4: 2回目に D \flat を打鍵したときの $Pkey_{k,2}$ (%)

調	C	D \flat	D	E \flat	E	F	G \flat	G	A \flat	A	B \flat	B
$Pkey_{k,1}(1)$	5.5	3.4	1.3	6.5	3.4	6.9	3.4	2.9	54.8	3.4	5.2	3.4
$Pkey_{k,2}^{\triangleright}(2)$	2.7	3.4	1.3	3.3	3.4	3.4	3.4	1.5	54.8	3.4	2.6	3.4
$count_{k,2}(3)$	0	2	1	1	2	0	2	0	2	2	0	2
$Pkey_{k,2}', Pkey_{k,2}(4)$	2.8	5.3	2.2	4.2	5.3	3.4	5.3	1.5	56.7	5.3	2.6	5.3

3.3.2 推定調をもとにした打鍵時のコード推定手法

以下ではコードを (a, b) と表現する。 a はコードのルート音 (コードの基準音) を示し、 $a = C, Db, D, Eb, E, F, Gb, Ab, A, Bb, B$ の 12 種類である。 b はコードの属性を示し、 $M7, m7, 7, mM7, sus4, dim, m7-5$ の 7 種類とする。時刻 t_n における、すなわち n 回目の打鍵を含む小節は $n - 1$ 回目打鍵や $n + 1$ 回目打鍵も含むことがある。ここではコードは小節ごとに設定するが、便宜的にそれぞれ $code_{n-1}, code_n, code_{n+1}$ と表記し、また同じ小節内の異なる打鍵番号により示されるコードは同一のものとする。なお、小節の区切りを自動的に認識するためにはテンポや拍子を推定する必要があるが、ここではテンポと拍子は固定して扱う。コード $code_n$ は、コード (a, b) ごとの確率 $P_{code(a,b),n}$ にもとづき推定する。コード $code_n$ は、時刻 t_n においてコード (a, b) であると判断する確率 (コード確率) $P_{code(a,b),n}$ のうち、最も確率が高いコード (a, b) とする。時刻 t_n においてコード (a, b) である確率は以下の 3 つの要素をもとに計算する。

1 つ目は、仮打鍵音 S_{pos_n} がコード (a, b) のコードトーン (付録 A.3) に含まれない場合は等しく低くなるような、またコードトーンに含まれる場合には等しく高くなるような確率 $P_{code-tone(a,b),n}$ を考える。例えば仮打鍵音 S_{pos_n} が C だった場合、コードトーンに C が含まれるコードは表 3.5 の 27 種類である。これらの各コード確率 $P_{code-tone(a,b),n}$ は同じ値 ($100/27=3.7\%$) とし、コードトーンでない各コード確率 $P_{code-tone(a,b),n}$ は 0 とする。

2 つ目はジャズバイブルのサンプル 71 曲の中でコード (a, b) が使用されている確率 $P_{code-ratio(a,b)}$ を考える。表 3.6 はこのサンプルの統計であるが、各行は調におけるその基準音とコードのルート音 a が何度の関係にあるか示している。すなわち、コードとは絶対的な音名 a とコードの属性 b で表すこともできるが、調における基準音に対する相対的な音程 (階名) a' とコードの属性 b で表すこともできる。例えば、調が C メジャーであるとき、コード CM7 は 1 度 (I) のメジャーセブンス (M7) として集計するが、調が G メジャーであるときは、CM7 は 4 度 (IV) のメジャーセブンス (M7) として集計し、また調が G メジャーであるときの GM7 は 1 度 (I) のメジャーセブンス (M7) として集計する。この表 3.6 は、絶対的な音名に対するコードの出現

割合をまとめたものではなく、各調における相対的な階名に対するコードの出現割合をまとめたものである。その上で調が推定されているときには、その調と各コードのルート音 a' をもとに、コード (a, b) の出現が期待できる確率 $P_{code-ratio(a,b)}$ を表 3.6 から取得する。

3つ目はサンプルデータを利用した確率 $P_{code-combi(a,b),n}$ である。これは時刻 t_n に対して、サンプルデータの3通りの集計を使い分けて決定する。

まず時刻 t_1 において、1回目の打鍵を行った時点では、既に奏でられた打鍵音推移はなく、仮打鍵音 S_{pos_1} の1音だけでユーザのイメージするコードの推定を行うことは非常に困難である。そこで同じくジャズバイブルのサンプル71曲を、調 k ごとに第一音 s で分類し、さらに第一音 s ごとに該当曲最初のコード (a', b) を集計する(表 3.7)。この表をもとに、第一音(仮打鍵音 S_{pos_1}) s 、推定調 k であるときにコード (a, b) である確率 $P_{code-combi-first(a,b),k,s}$ を求める。なお、サンプルデータ上で、曲の一番最初に用いられるコードは IM7, IIm7, II7, IIIIm7, IVM7, IV#m7-5, VIbM7, VIIm7, VIIIm7 の9種類であった。表 3.7 は必要な箇所のみを抜粋した表である。その上で、時刻 t_1 のときの各コードごとの確率 $P_{code-combi(a,b),n}$ は式 3.7 とする。

$$P_{code-combi(a,b),n} = P_{code-combi-first(a,b),k,s} \quad (3.3)$$

続いて、 $n \geq 2$ の時刻 t_n において、 $n-1$ 回目の打鍵と n 回目の打鍵が異なる小節で打鍵された場合、現在の小節に対応するコード (a, b) を決定するための情報として、前の小節で用いられたコード (α, β) を利用する。楽曲において、コードの進行はある程度定まっているため、音楽の知識がある者であれば、これまでのコードの推移から次のコードを予測することが可能である。そこで本研究では次のコードを推定するために、サンプルデータの前の小節のコードが (α, β) である場合に現在のコードが (a, b) である割合を調べた。このデータセットにより決定するコード (a, b) ごとの確率を $P_{code-combi-next(a,b),(\alpha,\beta),k}$ とする。ルート音12種類、コードの属性7種類、全部で84種類のコードが存在する。よって、 $P_{code-combi-next(a,b),(\alpha,\beta),k}$ では84種類×84種類=7056通りのコード進行が考慮されている。以下に $P_{code-combi-next(a,b),(\alpha,\beta),k}$

で用いられるデータセットから、コード (α, β) がIM7であるとき、現在の小節のコードが (a, b) である確率を抜粋した表3.8を示す。なおこれも、音名ではなくそれぞれの調における階名により集計している。その上で、時刻 t_n と時刻 t_{n-1} が異なる小節で打鍵された場合の各コード確率 $P_{code-combi(a,b),n}$ は式3.4とする。

$$P_{code-combi(a,b),n} = P_{code-combi-next(a,b),(\alpha,\beta),k} \quad (3.4)$$

最後に、同じく $n \geq 2$ の時刻 t_n であるが、 $n - 1$ 回目の打鍵と n 回目の打鍵が同じ小節の場合には、 P_{code_n} は単純に次式3.5により決定する。

$$P_{code-combi(a,b),n} = P_{code-combi(a,b),n-1} \quad (3.5)$$

以上の手法により求めた $P_{code-tone(a,b)}$, $P_{code-ratio(a,b)}$, $P_{code-combi(a,b),n}$ を用いて時刻 t_n におけるコード確率 $P_{code(a,b),n}$ を式3.6によって求める。また、このときの W_{tone} , W_{ratio} , W_{combi} は加重平均を求める際の重みであり、 $W_{tone} + W_{ratio} + W_{combi} = 1$ である。

$$P_{code(a,b),n} = P_{code-tone(a,b)} \times W_{tone} + P_{code-ratio(a,b)} \times W_{ratio} + P_{code-combi(a,b),n} \times W_{combi} \quad (3.6)$$

これにより得られた $P_{code(a,b),n}$ のうち、最も確率が高いときの値を C_{val_n} とし、コード (a, b) を時刻 t_n のコード $code_n$ と決定する。

表 3.5: C の音が構成音として含まれるコード

CM7	Cm7	C7	CmM7	Csus4	Cdim	Cm7-5
DbM7	DbmM7	Dm7	D7	Dm7-5		
Ebdim						
FM7	Fm7	F7	FmM7	Fsus4		
Gbdim	Gbm7-5					
Gsus4						
AbM7	AbmM7					
Am7	Am7-5	Adim	Am7-5			

表 3.6: コード (a', b) が全サンプル曲で使用された割合 (%)

$a' \setminus b$	M7	m7	7	mM7	sus4	dim	m7-5
I	12.53	0.84	2.42	0.07	0.00	0.03	0.00
IIb	0.37	0.13	0.50	0.00	0.00	0.07	0.00
II	0.37	14.21	4.13	0.03	0.00	0.00	0.50
IIIb	0.34	0.60	0.94	0.00	0.00	1.14	0.00
III	0.24	5.85	4.13	0.07	0.00	0.00	1.18
IV	3.70	1.71	2.15	0.03	0.00	0.00	0.00
Vb	0.00	0.47	0.27	0.00	0.00	0.47	1.48
V	0.20	1.85	13.84	0.00	0.50	0.00	0.00
VIb	0.57	0.20	0.91	0.00	0.00	0.20	0.00
VI	0.20	6.38	6.45	0.00	0.00	0.00	0.03
VIIb	0.34	0.40	2.62	0.00	0.00	0.00	0.03
VII	0.07	0.24	1.48	0.00	0.00	0.00	2.49

表 3.8: 前の小節のコード (α, β) が IM7 であるとき, 現在の小節のコードが (a', b) である確率 (%)

ルート音 (a)	M7	m7	7	mM7	sus4	dim	m7-5
I	0	3.14	3.71	0	0	0	0
II \flat	0.57	0.29	0.57	0	0	0.57	0
II	0.29	14	2	0	0	0	0.29
III \flat	0.29	2.86	0	0	0	1.14	0
III	0	1.43	5.43	0	0	0	3.14
IV	6.86	2	9.14	0	0	0	0
V \flat	0	0.29	0.29	0	0	0	4
V	0	6	3.14	0	2.57	0	0
VI \flat	0.86	0	0	0	0	0	0
VI	0	8.57	5.14	0	0	0	0
VII \flat	0.29	0	1.14	0	0	0	0
VII	0	0.29	0.29	0	0	0	9.43

以下では節 3.3.1 と同様に 1 回目の打鍵が Ab, 2 回目の打鍵が Db であったときの例を示す. このとき, 曲の調は Ab であるとする. まず, 時刻 t_1 のときに Ab を打鍵した場合, Ab をコードトーンに含むコードは, 以下の 27 種類 (表 3.9) である. よってこれらのコード (a, b) におけるコード確率 $P_{code-tone(a,b),1}$ はそれぞれ $1/27 = 3.7$ とし, それ以外のコード確率 $P_{code-tone(a,b),1}$ は 0 とする (表 3.10). また, ジャズバイブルのサンプル 71 曲の中でコード (a, b) が使用されている確率 $P_{code-ratio(a,b)}$ は表 3.6 の通りである. サンプルデータを利用した確率 $P_{code-combi(a,b),1}$ は打鍵を行ったときの状況によって決定する. 1 回目の打鍵を行った場合には, $P_{code-combi(a,b),1}$ には $P_{code-combi-first(a,b),k,s}$ の値を用いる. 表 3.7 より, 演奏曲の調が Ab であるときに Ab が打鍵された場合, 時刻 t_1 におけるコードは Fm7 である確率が 100% である. よって, コードが Fm7 に対する $P_{code-combi(a,b),1}$ は 100 とし, それ以外のコード (a, b) の確率は 0 とする. これらの $P_{code-tone(a,b),1}$, $P_{code-ratio(a,b)}$, $P_{code-combi(a,b),1}$ より, 式 3.11 を用いてコード確率 $P_{code(a,b),1}$ を決定する (表 3.11). なお, このときの重みは, $W_{tone} : W_{ratio} : W_{data} = 1 : 2 : 7$ とする. 以上により, 時刻 t_1 のコードは Fm7 と決定する.

次に, 時刻 t_2 に Db を打鍵した場合, Db をコードトーンに含むコードは, 以下の 27 種類 (表 3.12) である. よってこれらのコード (a, b) におけるコード確率 $P_{code-tone(a,b),2}$ はそれぞれ $1/27 = 3.7$ とし, それ以外のコード確率 $P_{code-tone(a,b),2}$ は 0 とする (表 3.13). また, 先ほどと同様, 確率 $P_{code-ratio(a,b)}$ は表 3.6 の通りである. サンプルデータを利用した確率 $P_{code-combi(a,b),1}$ は打鍵を行ったときの状況によって決定する. 時刻 t_2 が時刻 t_1 と同じ小節内であった場合, $P_{code-combi(a,b),2}$ では $P_{code-combi(a,b),1}$ を用いる. 一方で, 時刻 t_2 が時刻 t_1 と異なる小節であった場合, $P_{code-combi(a,b),2}$ には $P_{code-combi-next(a,b),(Fm7),Ab,2}$ の値を用いる (表 3.14). これらの $P_{code-tone(a,b),2}$, $P_{code-ratio(a,b)}$, $P_{code-combi(a,b),2}$ より, 式 3.11 を用いてコード確率 $P_{code(a,b),2}$ を決定する. 時刻 t_2 が時刻 t_1 と同じ小節内である場合の結果を表 3.15 に示し, 時刻 t_2 が時刻 t_1 と異なる小節内である場合の結果を表 3.16 に示す. 以上より, 時刻 t_2 が時刻 t_1 と同じ小節内である場合の推定コード $code_2$ は Fm7, 異なる小節の場合は Bb7 とする.

表 3.9: Ab の音が構成音に含まれるコード

D♭M7	D♭m7	D♭7	D♭mM7	D♭sus4		
Ddim	Dm7-5					
E♭sus4						
EM7	E7					
Fm7	FmM7	Fdim	Fm7-5			
A♭M7	A♭m7	A♭7	A♭mM7	A♭sus4	A♭dim	A♭m7-5
AM7	AmM7					
B♭m7	B♭7	B♭m7-5				
Bdim						

表 3.10: 1 回目に Ab を打鍵した場合の確率 $P_{code-tone_{(a,b),1}}(\%)$

ルート音 (a)	M7	m7	7	mM7	sus4	dim	m7-5
C	0	0	0	0	0	0	0
D♭	3.7	3.7	3.7	3.7	3.7	0	0
D	0	0	0	0	0	3.7	3.7
E♭	0	0	0	0	3.7	0	0
E	3.7	0	3.7	0	0	0	0
F	0	3.7	0	3.7	0	3.7	3.7
G♭	0	0	0	0	0	0	0
G	0	0	0	0	0	0	0
A♭	3.7	3.7	3.7	3.7	3.7	3.7	3.7
A	3.7	0	0	3.7	0	0	0
B♭	0	3.7	3.7	0	0	3.7	0
B	0	0	0	0	0	3.7	0

表 3.11: 1 回目に A_b を打鍵した場合のコード確率 $P_{code(a,b),1}(\%)$

ルート音 (a)	M7	m7	7	mM7	sus4	dim	m7-5
C	0.04	1.17	0.83	0.01	0	0	0.24
D_b	1.11	0.71	0.80	0.38	0.37	0	0
D	0	0.09	0.05	0	0	0.46	0.67
E_b	0.04	0.37	2.77	0	0.47	0	0
E	0.48	0.04	0.55	0	0	0.04	0
F	0.04	71.64	1.29	0.37	0	0.37	0.37
G_b	0.06	0.08	0.52	0	0	0	0.01
G	0.01	0.04	0.30	0	0	0	0.50
A_b	2.88	0.53	0.38	0.38	0.37	0.38	0.37
A	0.44	0.02	0.10	0.37	0	0.01	0
B_b	0.07	3.21	1.20	0.01	0	0	0.47
B	0.07	0.12	0.19	0	0	0.60	0

表 3.12: D_b の音が構成音に含まれるコード

D_bM7	D_bm7	D_b7	D_bmM7	D_bsus4	D_bdim	D_bm7-5
DM7	DmM7					
E_bm7	E_b7	E_bm7-5				
Edim						
G_bM7	G_bm7	G_b7	G_bmM7	G_bsus4		
Gdim	$Gm7-5$					
A_bsus4						
AM7	A7					
B_bm7	B_bmM7	B_bdim	B_bm7-5			

表 3.13: 2 回目に Db を打鍵した場合の確率 $P_{code-tone_{(a,b),1}}(\%)$

ルート音 (a)	M7	m7	7	mM7	sus4	dim	m7-5
C	0	0	0	0	0	0	0
Db	3.7	3.7	3.7	3.7	3.7	3.7	3.7
D	3.7	0	0	3.7	0	0	0
Eb	0	3.7	3.7	0	0	3.7	0
E	0	0	0	0	0	3.7	0
F	0	0	0	0	0	0	0
Gb	3.7	3.7	3.7	3.7	3.7	0	0
G	0	0	0	0	0	3.7	3.7
Ab	0	0	0	0	3.7	0	0
A	3.7	0	3.7	0	0	0	0
Bb	0	3.7	0	3.7	0	3.7	3.7
B	0	0	0	0	0	0	0

表 3.14: 1 小節前のコードが Fm7 であるときの $P_{code-combi-next_{(a,b),(Fm7),Ab,2}}(\%)$

ルート音 (a)	M7	m7	7	mM7	sus4	dim	m7-5
C	0	1.05	4.21	0	0	0	0
Db	3.68	0.52	1.05	0	0	0	0
D	0	0.52	0	0	0	0	6.84
Eb	0	2.63	0.52	0	0	0	0
E	0	0	3.15	0	0	0	0
F	0	0	2.10	0	0	0	0
Gb	0.53	0.53	0	0	0	0	0
G	0	0	0	0.53	0	0	5.26
Ab	0.53	0	0.53	0	0	0	0
A	0	0	0	0	0	0	0
Bb	1.58	23.68	37.89	0	0	0	0
B	0	0.53	2.10	0	0	0	0

表 3.15: 時刻 t_2 が時刻 t_1 と同じ小節内である場合のコード確率 $P_{code(a,b),2}(\%)$

ルート音 (a)	M7	m7	7	mM7	sus4	dim	m7-5
C	0.08	1.99	1.40	0.02	0	0	0.40
D \flat	1.89	1.21	1.36	0.64	0.63	0.37	0.37
D	0.37	0.15	0.09	0.37	0	0.42	0.76
E \flat	0.06	1.00	5.08	0	0.43	0	0.37
E	0.45	0.07	0.57	0	0	0.44	0
F	0.07	51.43	2.19	0.26	0	0.26	0.27
G \flat	0.49	0.50	1.26	0.37	0.37	0	0.01
G	0.02	0.08	0.50	0	0	0.37	1.21
A \flat	4.52	0.54	1.08	0.28	0.63	0.27	0.26
A	0.76	0.04	0.54	0.26	0	0.02	0
B \flat	0.13	5.46	1.66	0.38	0	0.37	0.80
B	0.12	0.20	0.32	0	0	0.65	0

表 3.16: 時刻 t_2 が時刻 t_1 と異なる小節内である場合のコード確率 $P_{code(a,b),2}(\%)$

ルート音 (a)	M7	m7	7	mM7	sus4	dim	m7-5
C	0.05	1.91	3.77	0.01	0	0	0.24
D \flat	3.69	1.08	1.54	0.38	0.38	0.38	0.38
D	0.38	0.46	0.05	0.37	0	0.09	5.09
E \flat	0.04	2.58	3.51	0	0.10	0	0.37
E	0.11	0.04	2.39	0	0	0.41	0
F	0.04	1.28	2.76	0	0	0	0.01
G \flat	0.80	0.81	0.89	0.37	0.37	0	0.01
G	0.01	0.05	0.66	0	0	0.37	4.55
A \flat	2.88	0.17	0.85	0.01	0.37	0.01	0
A	0.44	0.03	0.47	0	0	0.01	0
B \flat	1.17	19.79	27.35	0.37	0	0.37	0.47
B	0.07	0.49	1.66	0	0	0.23	0

3.3.3 推定調と推定コードを用いた打鍵確率推定

節 3.3.1 で決定した推定調 key_n と 3.3.2 で決定した推定コード $code_n$ を用いて、時刻 t_n の音楽理論をもとにした各鍵盤の打鍵確率 (理論打鍵確率) $Pthe_{S,n}$ を決定する。まず、演奏曲の調が key_n であるときの各鍵盤の打鍵確率 $Pthe-key_{S,n}$ を求める。このとき、打鍵確率を実際に求める対象の鍵盤は、3.2 節の二等辺三角形の範囲に含まれる鍵盤である。ここで、調 k のダイアトニックスケールの構成音である二等辺三角形内の鍵盤の数を $count-key_{k,n}$ とする。鍵盤 S が調 $k = key_n$ のダイアトニックスケールだった場合、鍵盤の打鍵確率 $Pthe-key_{S,n} = 1/count-key_{k,n}$ とし、ダイアトニックスケールでなかった場合には $Pthe-key_{S,n} = 0$ とする。

次にコード $code_n$ における各鍵盤の打鍵確率 $Pthe-code_{S,n}$ を求める。 $Pthe-key_{S,n}$ を決定する時と同様に 3.2 の二等辺三角形の範囲内の鍵盤で、コード c のコードトーンに含まれる鍵盤の数を $count-code_{c,n}$ とする。鍵盤 S がコード $c = code_n$ のコードトーンだった場合、鍵盤の打鍵確率 $Pthe-code_{S,n} = 1/count-code_{c,n}$ とし、コードトーンでなかった場合には $Pthe-code_{S,n} = 0$ とする。最後に $Pthe-key_{S,n}$ と $Pthe-code_{S,n}$ を式 3.7 のように加重平均をとることにより、音楽理論をもとにした各鍵盤の打鍵確率 $Pthe_{S,n}$ を求める。また、式 3.7 で用いられる、 $Wthe-key, Wthe-code$ は加重平均を取る際の重みであり、 $Wthe-key + Wthe-code = 1$ である。

$$Pthe_{S,n} = Pthe-key_{S,n} \times Wthe-key + Pthe-code_{S,n} \times Wthe-code \quad (3.7)$$

例えば、図 3.4 において、時刻 t_n での調 key_n が Bb 、コード $code_n$ が $Cm7$ だった場合を想定する。 Bb メジャーキーのダイアトニックスケールは Bb, C, D, Eb, F, G, A によって構成され、 $Cm7$ は C, Eb, G, Bb によって構成される。図 3.4 (1) に示す赤い鍵盤の箇所が調 key_n のダイアトニックスケールの構成音に当たるため、 $count-key_{k,n} = 8$ となる。よって調 key_n における各鍵盤の打鍵確率は $Pthe-key_{S,n} = 1/8$ ($S = D, Eb, F, G, A, Bb, C, D$) のように決定する。同様に図 3.4 (2) に示す赤い鍵盤の箇所がコード $code_n$ のコードトーンに当たるため、 $count-code_{c,n} = 4$ となる。よってコード $code_n$ における各鍵盤の打鍵確率は $Pthe-code_{S,n} = 1/4$ ($S = Eb, G, Bb, C$) と

求めることが可能である. Wthe-key と Wthe-code をそれぞれ0.5とした場合の結果を表3.17に示す.

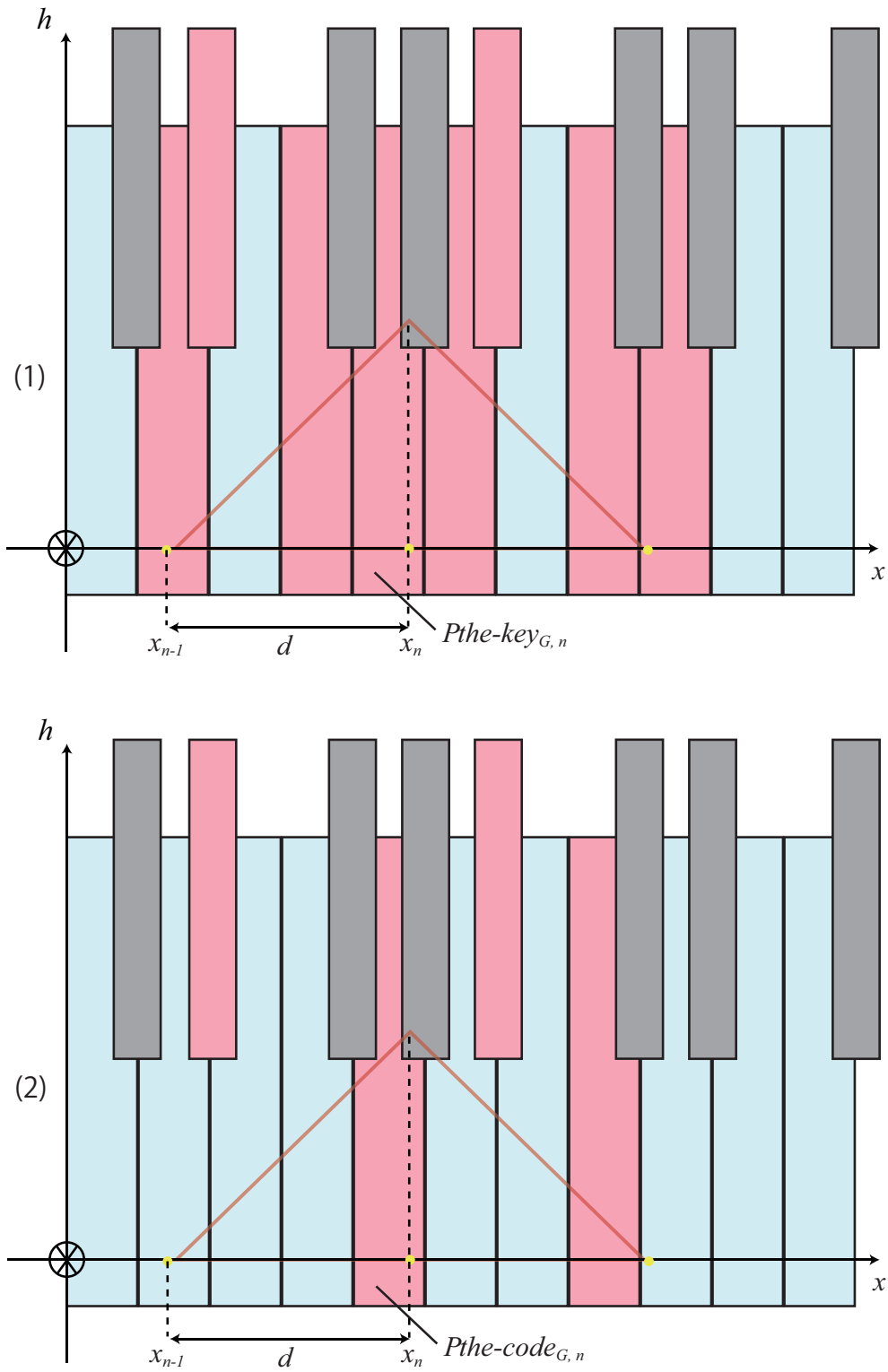


図 3.4: $key_n=Bb$, $code_n=Cm7$ のときの $Pthe-keys_n$ と $Pthe-codes_n$

表 3.17: 図 3.4 における鍵盤 S ごとの打鍵確率 (%)

鍵盤 S	$P_{the-keys,n}$	$P_{the-codes,n}$	$P_{the-S,n}$ ($W_{the-key} : W_{the-code} = 0.5 : 0.5$)
D	1/8	0	1/16
E \flat	1/8	1/4	3/16
E	0	0	0
F	1/8	0	1/16
G \flat	0	0	0
G	1/8	1/4	3/16
A \flat	0	0	0
A	1/8	0	1/16
B \flat	1/8	1/4	3/16
B	0	0	0
C	1/8	1/4	3/16
D \flat	0	0	0
D	1/8	0	1/16

3.4 位置打鍵確率と理論打鍵確率を用いた打鍵音確率推定

最終的な打鍵音は打鍵位置による推定手法と音楽理論による推定手法を統合した上で推定、決定する。なお、ユーザは手動で打鍵位置と音楽理論の信頼度を設定することができる。例えばユーザが音楽経験者であり、打鍵した位置に自信を持って演奏していた場合には、音楽理論の信頼度を低く設定し、打鍵位置の信頼度を高く設定することでユーザの思い描いたメロディに近い演奏を実現する。反対に、ユーザが音楽未経験者であり、演奏を行ったとしても自分の思い通りの鍵盤を押す自信がない場合や、そもそも適切なメロディを思い浮かべられない場合には、音楽理論の信頼度を高く設定し、打鍵位置の信頼度を低く設定することで、おおよそランダムな打鍵に対しても違和感のないメロディを実現する。

しかし、音楽理論の信頼度とはいうが、そもそも推定した調やコード自体の信頼度が完全ではない。なぜなら、調においてもコードにおいても最も確率が高かった調やコードを選択して採用しているだけであり、採用されなかった調やコードの確率も採用されたものの確率を超えない程度に高い場合も存在するためである。例えば、調Cメジャーの採用確率 $Pkey_{C,n}$ が30%、Fメジャーの採用確率 $Pkey_{F,n}$ が20%、その他の採用確率 $Pkey_{k,n}$ が5%ずつだったとする。時刻 t_n における調がCメジャーである確率が30%しかなく、他の調である確率が70%もある中でこれを音楽理論として完全に信頼して最終的な打鍵音を決定することはあまり妥当とは言えない。また、これはコードにおいても同様のことが言える。

そのため、本システムではユーザが決定した音楽理論の信頼度（指定信頼度）と、調とコードの信頼度から得られた音楽理論の信頼度（推定信頼度）の比によって最終的な音楽理論の信頼度を決定する。まず、ユーザが設定した打鍵位置の信頼度を $W_{pos-user}$ とし、音楽理論の理論の信頼度を $W_{the-user}$ とする ($W_{pos-user} + W_{the-user} = 1$)。また3.3.1で求めた調ごとの採用確率 $Pkey_{k,n}$ の中で最も高い値である $Kval_n$ をもとに調の信頼度を決定し、3.3.2で求めたコードごとの採用確率 $Pcode_{(a,b),n}$ の中で最も高かった値である $Cval_n$ を用いてコードの信頼度を決定する。これらの値から、時刻 t_n における実際の音楽理論信頼度 W_{the_n} と打鍵位置信頼度 W_{pos_n} ($W_{the_n} + W_{pos_n} = 1$) を決定した上で、理論打鍵確率と位置打鍵確率をも

とに最終的な打鍵確率を求める。

具体的には、いくつかのサンプル曲をもとに演奏実験を行った上で、経験的に以下のように重みを定義する。

$$Wthe_n = \frac{\left(-\frac{1}{100}Kval_n^2 + 2Kval_n\right) + \left(\frac{100}{3} \log 10 \times Cval_n + 1\right)}{2} \times Wthe-user$$

$$Wpos_n = 1 - Wthe_n \quad (3.8)$$

なお、音楽理論信頼度 $Wthe_n$ を算出するための各項の推移を図 3.5 に示す。その上で、時刻 t_n における鍵盤 S に対する打鍵確率 $P_{S,n}$ は次式により決定する。

$$P_{S,n} = Pthe_{S,n} \times Wthe_n + Ppos_{S,n} \times Wpos_n \quad (3.9)$$

打鍵確率 $P_{S,n}$ が最も高い S を時刻 t_n の打鍵音として決定し、奏でる。指定信頼度の設定によっては打鍵位置から離れた鍵盤が選択される場合がある。打鍵位置の鍵盤(音)と決定した打鍵音が3度以上離れている場合、打鍵位置と決定打鍵音の中心が一致するように鍵盤座標系を x 軸に平行にシフトさせる。なお、差が2度以内の場合には鍵盤座標はシフトさせない。

ここで時刻 t_1 の時に Ab を打鍵し、時刻 t_2 の時に時刻 t_1 とは異なる小節で Db を打鍵した場合の時刻 t_2 における打鍵音決定の例を示す。このときの指定信頼度は、 $Wpos-user : Wthe-user = 0.6 : 0.4$ とする。時刻 t_1 と時刻 t_2 の打鍵位置 (x_1, y_1) と (x_2, y_2) は図 3.6 の位置であるとする。時刻 t_2 において黒鍵を打鍵しているため、仮打鍵音決定のためのボーダーは黒鍵の幅を少し広げた幅で設定する。このときの $Ppos_{S,n}$ を表 3.18 (1) とする。節 3.3.1 と節 3.3.2 より、 $Pkey_{Ab,2} = 56.7\%$ であることより推定調は Ab であり $Kval_2 = 56.7$ である。同様に $Pcode_{Bb7,2} = 27.35\%$ であることより推定コードは $Bb7$ であり、 $Cval_2 = 27.35$ である。それに伴い、 $Pthe-keys_2$ (表 3.18 (2)), $Pthe-key_{S,2}$ (表 3.18 (3)), $Pthe_{S,2}$ (表 3.18 (4)) を決定する。なお、図 3.6(1) の桃色と紫色の箇所が二等辺三角形中の Ab メジャースケールのダイアトニックスケールの構成音である鍵盤であり、青色と紫色の箇所が $Bb7$ のコードトーンである鍵盤である。紫色の箇所はどちらにも含まれる鍵盤である。ここで $Wthe_2$ を求めるため、図 3.5 のグラフにより計算を行うと $-\frac{1}{100}56.7^2 + 2 \times 56.7 = 87.25$, $\frac{100}{3} \log 10 \times 27.35 + 1 = 81.28$ となり、 $Wthe-user = 0.4$ であることから $Wthe_2 =$

表 3.18: $W_{pos-user} : W_{the-user} = 0.6 : 0.4$, $S_1=Ab$, $S_2=Db$ のときの時刻 t_2 における打鍵確率 (%)

	Ab	A	Bb	B	C	Db	D	Eb	E	F	Gb
$P_{pos_{S,2}}(1)$	2.42	0.96	12.3	10.24	12.8	27.66	3.2	18.9	6.4	3.84	1.28
$P_{the-key_{S,2}}(2)$	16.66	0	16.66	0	16.66	16.66	0	16.66	0	16.66	0
$P_{the-code_{S,2}}(3)$	25	0	25	0	0	0	25	0	0	25	0
$P_{the_{S,2}}(4)$	20.83	0	20.83	0	8.33	8.33	12.5	8.33	0	20.83	0
$P_{S,2}(5)$	8.63	0.64	15.17	6.79	11.29	21.14	6.33	15.33	4.24	9.57	0.85
$P_{S,2}(6)$	14.83	0.31	18.05	3.33	9.78	14.63	9.46	11.77	2.08	15.29	0.41

33.71である。これより、 $W_{pos_2} = 66.29$ である。式 3.9 を用いて各鍵盤の打鍵確率 $P_{S,2}$ (表 3.18 (5)) を求める。以上により、打鍵音は Db に決定する (図 3.6 (2))。一方で $W_{pos-user} : W_{the-user} = 0.2 : 0.8$ のときの結果を表 3.18 (6) に示す。このときは、時刻打鍵音は Bb に決定する。打鍵位置から 3 度以上離れた鍵盤を打鍵音として決定し、鳴らすため、座標系をシフトさせ、次回の打鍵位置はこのシフトされた座標系において扱う。

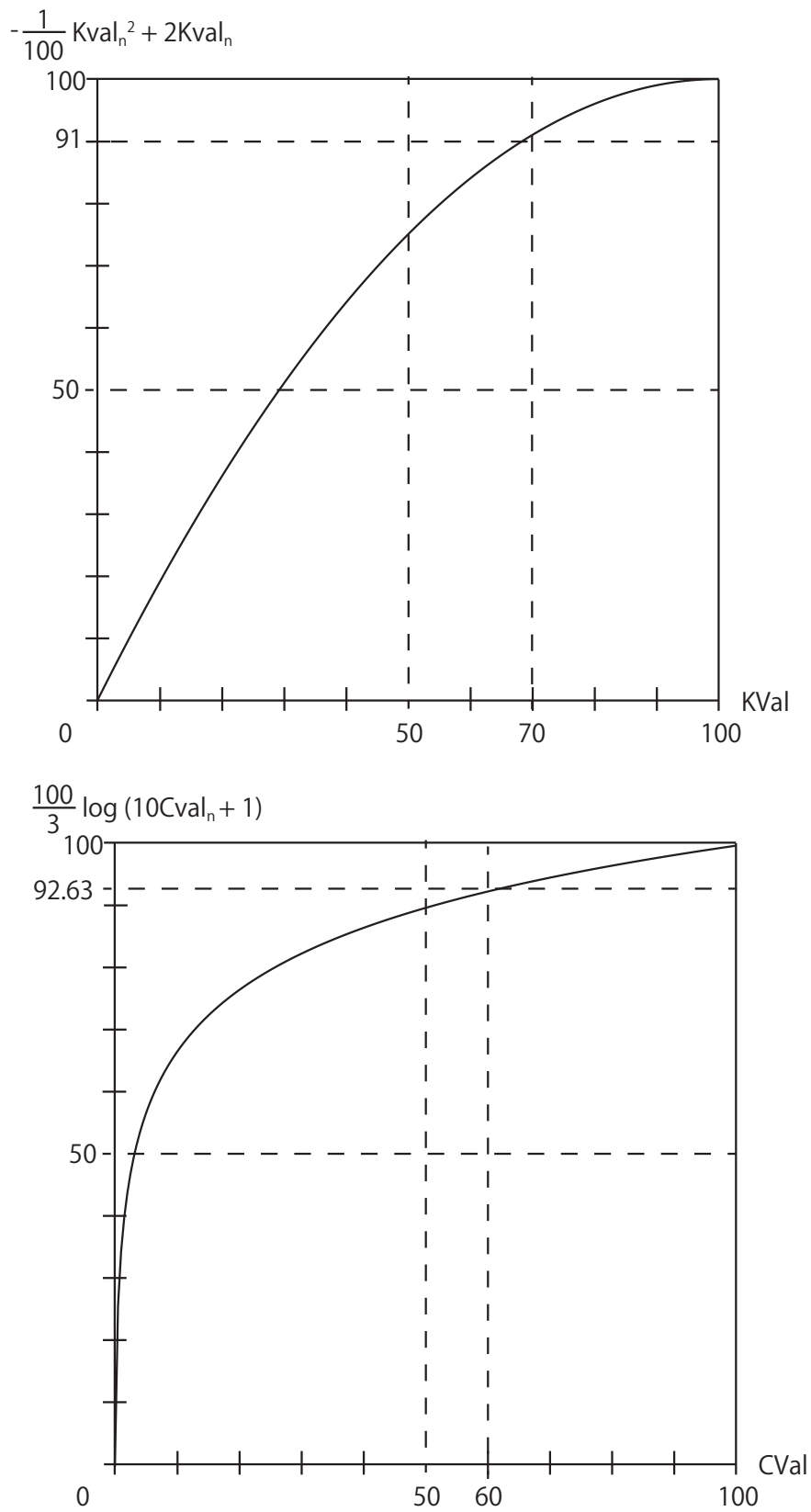


図 3.5: $-\frac{1}{100} Kval_n^2 + 2Kval_n$, $\frac{100}{3} \log 10 \times Cval_n + 1$ の推移

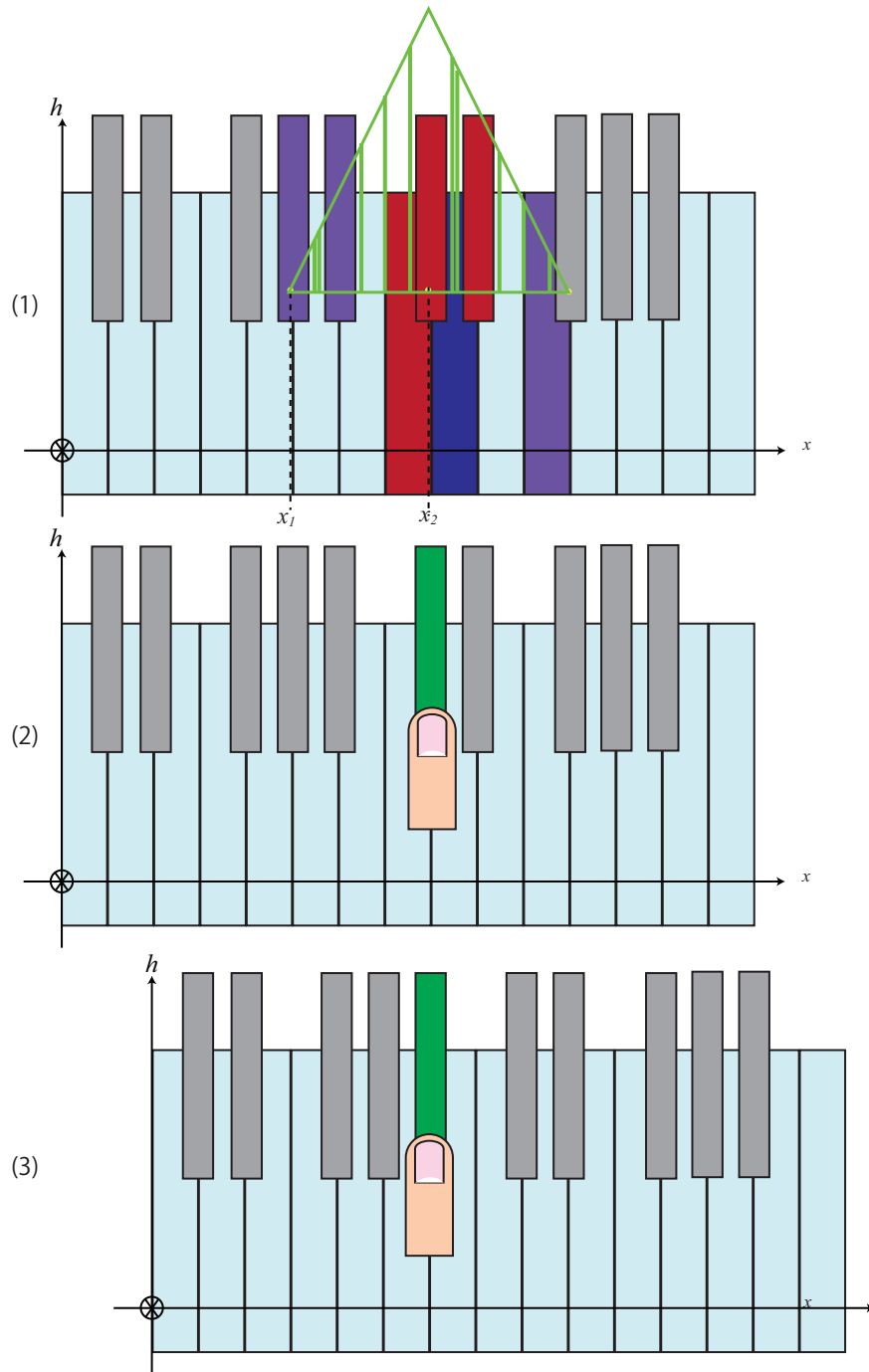


図 3.6: $S_1=Ab$, $S_2=Db$ のときの時刻 t_2 における打鍵確率 (%)

第4章 実験

4.1 実験環境

実験は 3.10GHz Intel Core i5-2400CPU とメモリ 8GB の環境で行い、動作確認のための画面表示には OpenCV と OpenGL を利用した。また、音声の生成には OpenAL を利用した。手指の取得には Leap Motion 社の Leap Motion を使用した。Leap Motion[25] は 2012 年にアメリカの Leap Motion 社から開発された、手のジェスチャーによってコンピューターの操作を可能とする入力機器である。マウスや画面タッチを用いず、ジェスチャーによって直感的な操作を可能にする。2基の赤外線カメラと赤外線照射 LED から構成されており、赤外線 LED に照らされた手や指を2基の赤外線カメラで撮影し、画像解析によって 3D 空間での手指の位置を推定することができる。半径 50cm 程度、中心角 110 度の範囲を検知することができ、手や指を 0.01 ミリの精度で認識することが可能である。本研究では Leap Motion を効果的に利用するために API[26] を利用する。

図 4.1 に示すように、Leap Motion の上にアクリル板を設置し、その上で指を動かすことによって VR ピアノを演奏する。ユーザが画面や手元を見ることなく演奏することを想定しているが、手のイニシャルポジションの確認のため、画面上にキーボードの CG を表示している。

テンポを固定するために、演奏時には常にメトロノームが流れており、ユーザはメトロノームに合わせて演奏を行う。また、演奏曲のコードは 1 小節ごとにのみ変化し、同じ小節内でコードが変わることはないものとする。

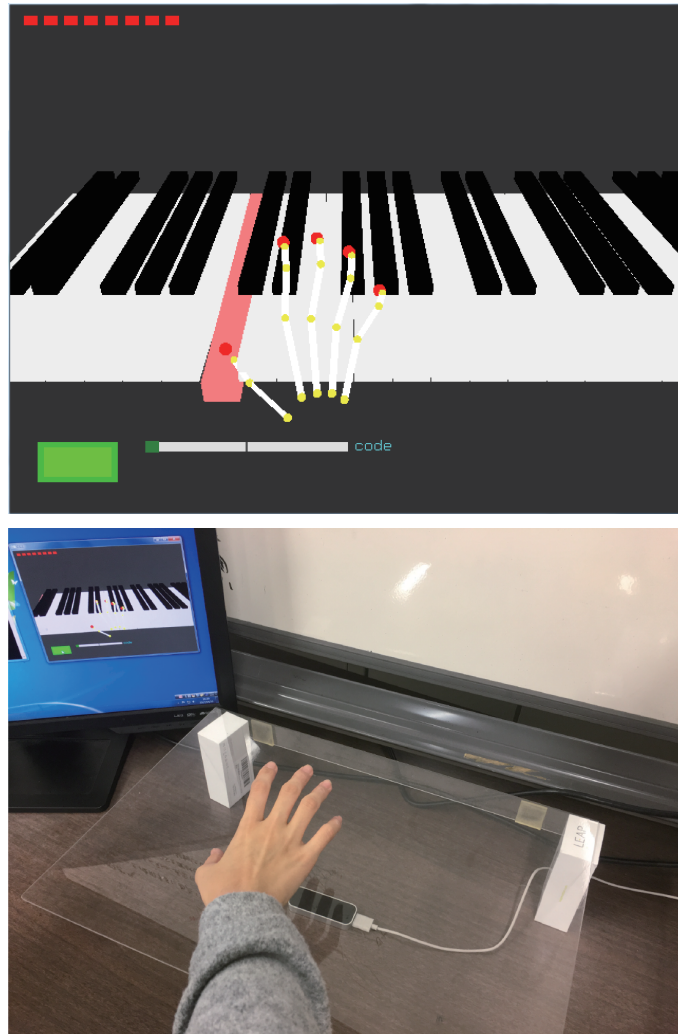


図 4.1: システムの外観

4.2 実験を行うにあたっての重み決定

3章で説明した打鍵音推定手法では、打鍵情報やサンプルをもとにしたデータを加重平均することによって調とコードの打鍵確率の決定を行っている。実験を行うにあたり、調の推定を行う時に用いる重み W_{key} , W_{data} , W_{first} と、コードの推定を行う時に用いる重み W_{tone} , W_{ratio} , W_{data} の最適値を求める必要がある。そこで、サンプルデータから無作為に選んだ3曲のデータを取得し、調推定のための重み W_{key} , W_{data} , W_{first} とコード推定のための重み W_{tone} , W_{ratio} , W_{data} の数値を決定する。

まず、サンプル曲3曲を用いて、調推定のための重みが $W_{key} + W_{data} + W_{first} = 1$ を維持するように、各値を0.1ずつ増減させ、4小節演奏し終わった後に推定された調について調べた。その後、3曲とも本来の調と同じ調であると判定した箇所を青色、それ以外を黄色となるように塗り分けを行ない、3曲の結果を統合した(図4.2)。なお図4.2では横軸が W_{key} 、縦軸が W_{data} の値を示す。図4.2により、本来の調と推定調が一致する可能性が高い、青色の範囲の中心である $W_{key} : W_{data} : W_{first} = 0.4 : 0.3 : 0.3$ を本実験で扱う調推定のための重みとして決定する。

同様にサンプル曲3曲を用いてコード推定のための重み W_{tone} , W_{ratio} , W_{combi} を決定する。なお、コードの種類は84種類であり、調に比べて数が多いことから、より多くのコード進行が結果として表示されることが予想されるためより細かくデータを調べる必要がある。よって、コードの推定のための重みを $W_{tone} + W_{ratio} + W_{combi} = 1$ を維持するように、0.05ずつ各値を増減させて4小節間演奏を行ったデータからコード進行を調べた。それぞれ4小節のコード進行が完全に一致した箇所ごとに色の塗り分けを行った結果(図4.3)、重みの比によって大きくA, B, Cの3パターンのコード進行が推定されることがわかった(図4.3)。なお図4.2では横軸が W_{combi} 、縦軸が W_{tone} の値を示す。サンプル曲1曲目の結果(図4.3(1))、サンプル曲2曲目の結果(図4.3(2))、サンプル曲3曲目の結果(図4.3(3))を統合した結果が図4.4であり、黒枠の範囲が3曲ともがコード進行パターンA, B, Cそれぞれの範囲である箇所を示す。これにより、コード進行パターンAを導き出す比率を $W_{tone} : W_{ratio} : W_{combi} = 0.1 : 0.4 : 0.5$ 、コード進行パターンBを導き出す比率

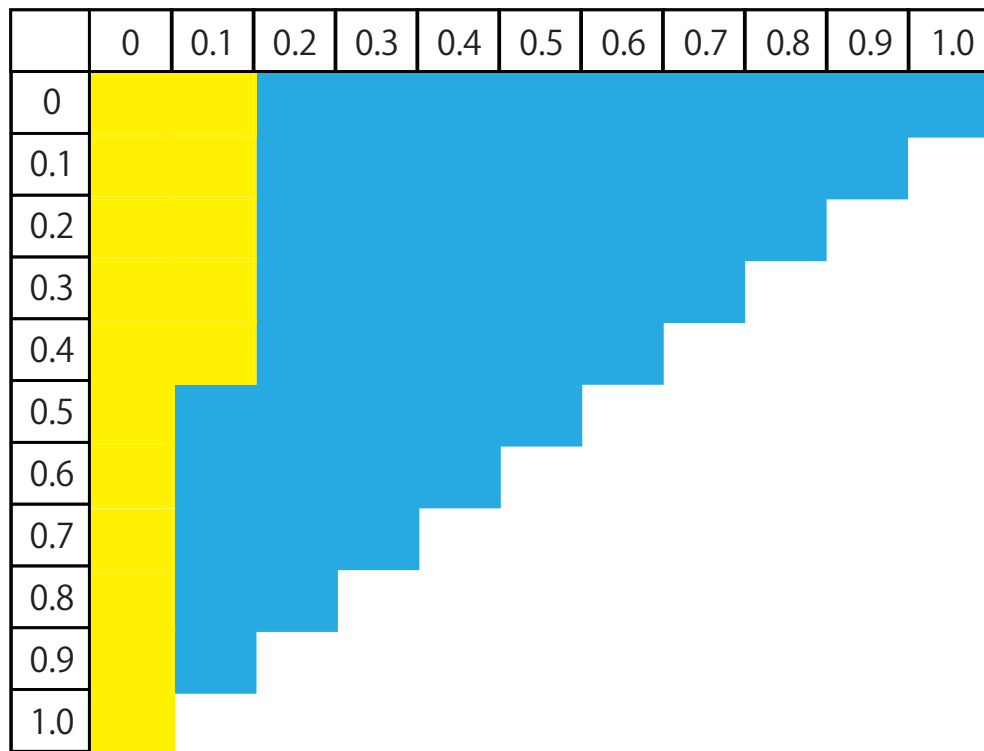


図 4.2: 調推定のための重みの推移

を $W_{tone} : W_{ratio} : W_{combi} = 0.3 : 0.3 : 0.4$, コード進行パターン C を導き出す比率を $W_{tone} : W_{ratio} : W_{combi} = 0.7 : 0.1 : 0.2$ と考えた。なお, コード進行パターン A, B, C は, A は打鍵した音に忠実ではあるが, 音楽の流れをあまり考慮しない傾向がある。B は本来のコード進行に最も近いコード進行が推定される傾向がある。C は本来のコード進行に近いものの, 一般的に頻出するコードが多く使われる傾向がある。よって, 実験を行うにあたりコード進行パターン B を用いる。コード進行パターン B を用いる妥当性を得るため, サンプル曲 3 曲とジャズバイブル以外の 3 曲に対して, コード進行パターン A, B, C により決定したコードを割り当てた音源 18 曲を音楽経験者, 音楽未経験者それぞれ 12 人ずつに聞いてもらい, 5 段階で評価をしてもらった。その結果, 音楽経験者によるコード進行パターン A, B, C の評価の平均点数は, A が 2.68 点, B が 3.73 点, C が 3.02 点であった。また, 音楽未経験者によるコード進行パターン A, B, C の評価の平均点数は, A が 2.98 点, B が 3.08 点, C が 3.14 点であった。本アンケートを行った音楽経験者には, ジャズのコード進行

に耳馴染みのある者が多かったため、もっともジャズらしいコード進行パターンであるBの点数が高くなった。一方で、音楽未経験者からはポップスでよく用いられるような比較的単調なコード進行が好まれ、Aのパターンの点数が最も高くなった。両者の点数の平均をとるとAが2.83点、Bが3.40点、Cが3.08点であることと、本論文で用いるサンプル曲71曲がジャズの曲であることから、コード進行パターンBである $W_{tone} : W_{ratio} : W_{combi} = 0.3 : 0.3 : 0.4$ を採用し、実験を行うことは妥当であると考えられる。

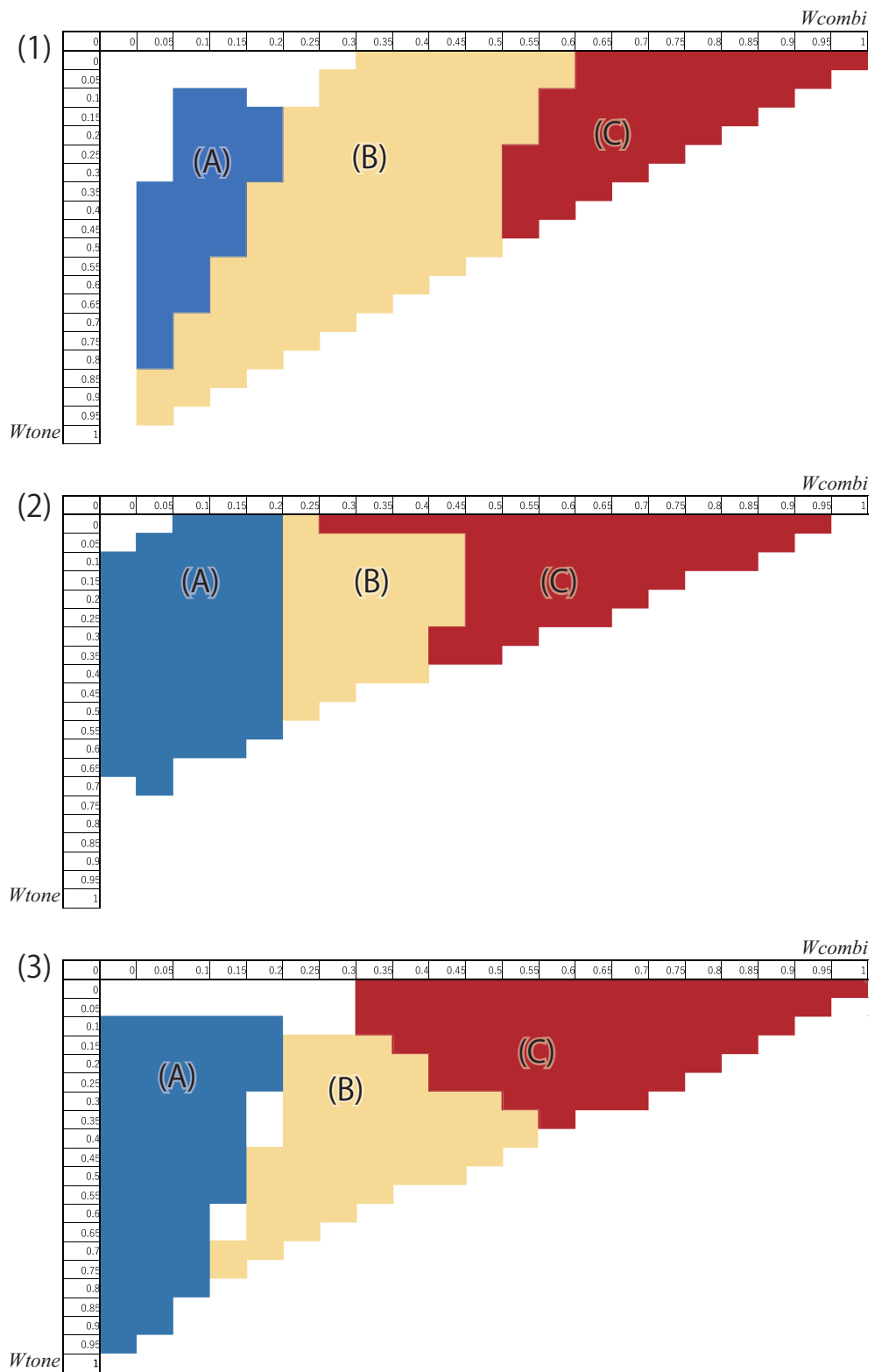


図 4.3: サンプルデータにおけるコード推定のための重みの推移

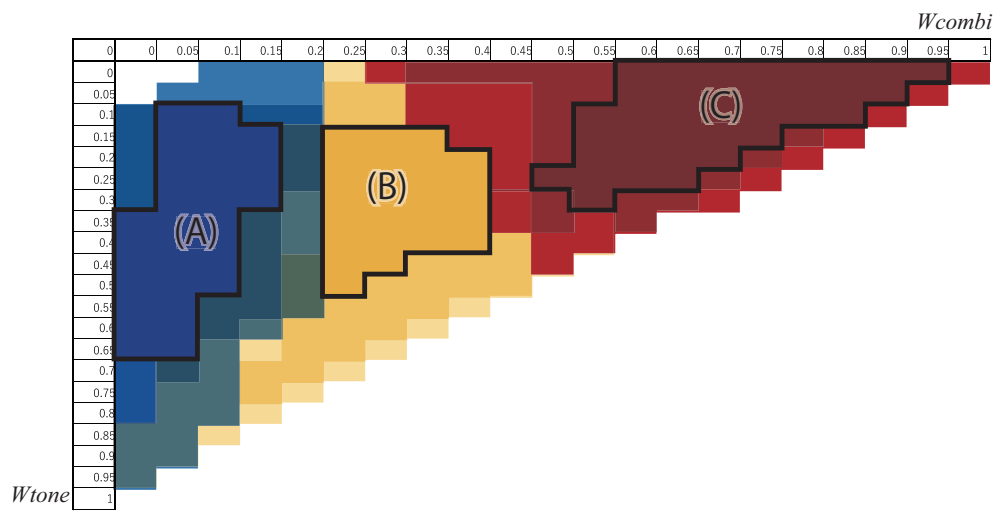


図 4.4: コード推定のための重みの推移

4.3 音楽経験者による演奏

提案手法によるシステムで、ユーザの意図していないミスをも、音楽理論により適切に修正できるかどうか評価する。そこで、音楽理論を理解した楽器経験者に確認用CG画面を見ずに4小節の曲を3曲演奏してもらい、図4.5に示すA,B,Cの運指データを取得した。実際の運指データはパターン1に相当する。演奏後に3種類の運指データに対し、表4.1の5パターンの指定信頼度を用いて必要に応じて補正を行う。なお、楽器経験者に対して用いる指定信頼度として想定しているパターンは1, 2, 3の3種類であるが、本実験では音楽未経験者を想定した指定信頼度のパターン4, 5も適用して演奏者によって自分の希望通りの演奏が行えたか評価を行う。なお、推定されたコード進行を、外部アプリケーション [27] によって再現し、メロディと合成することによって楽曲も作成する。それぞれの楽曲の音楽らしさを音楽経験者33人、音楽未経験者21人に評価してもらうことで、運指ミス修正の手がかりとしているコード進行推定の妥当性を評価した(表4.2)。それぞれの評価は、伴奏とメロディに違和感がない場合に5点、違和感がある場合に1点として、5点満点で行った。図4.5は各運指A, B, Cに対して5パターンの指定信頼度を用いて補正を行った結果と、その過程のコード進行である。図中の赤枠で囲まれた箇所は指定信頼度の比の違いにより、選択された打鍵音が変化した部分である。

楽器経験者による演奏は、無意識的にコード進行をイメージしながら行っていたため、推定調の曲の途中でのや、全てのパターンにおいてのコードの変化が少なかった。そのため、音楽理論の指定信頼度の比率を上げた場合でもシステムによる補正はあまりなく、安定した演奏となった。運指Aの演奏時において、演奏者は赤枠の音で誤打鍵をしており、希望通りの演奏ができなかった。システムによって補正を行った結果、誤打鍵をした運指Aにおいては、音楽経験者用の指定信頼度であるパターン2, 3を用いることで希望通りのメロディを出力できた。また、パターン4, 5においても同様の結果を得ており、本論文の提案手法の妥当性が示されている。アンケートによる結果(表4.2A(1), A(2))においても希望通りのメロディに補正が行われている図4.5のA(2)の方がA(1)よりも音楽的に優れているという結果を得ることができた。運指Bの演奏において、演奏者は希望通りの演奏ができた。そのデータ

に対し、システムによって音楽経験者用の信頼度であるパターン1, 2, 3を用いた結果、特に補正が行われることなく希望通りのメロディが出力された。一方で音楽未経験者用の指定信頼度であるパターン4, 5を用いた結果、演奏者の意思とは異なる補正が行われた。図4.5のB(1)とB(2)の楽曲におけるアンケート結果(表4.2B(1), B(2))においても演奏者の希望通りに奏でられたメロディの方が僅かではあるが優れているという結果が得られた。なお、パターン4, 5により補正が行われている楽曲においても音楽的に優れているという結果が得られており、補正結果は演奏者自身の意思とは異なるものの、本論文の提案手法自体の妥当性は示されている。運指Cの演奏においても、演奏者は希望通りの演奏ができた、そのデータに対し、システムによってパターン1, 2, 3, 4を用いた結果、特に補正が行われることなく希望通りのメロディが出力された。一方でパターン5を用いた結果、演奏者の意思とは異なる補正が行われた。アンケートによる結果においても希望通りのメロディに補正が行われている図4.5のC(1)がC(2)よりも音楽的に優れているという結果を得ることができた。しかし、パターン4, 5によって補正されたC(2)は音楽的に大きな違和感を生んでしまっており、システムで用いる音楽理論に問題点があることが示唆された。この違和感の原因は最初の1小節で用いられているBの音である。Dm7のコードに対して1音目のBの音はアボイドノート(付録A.4)と呼ばれ、経過音として少しだけ用いる分には問題ないが、強調して音を長く使ってしまうと大きな違和感を生んでしまう音である。C(2)ではこのアボイドノートを、2拍に渡って強調してしまっているため、大きな違和感を生み出しているものと思われる。本提案手法では音楽理論においてアボイドノートを考慮していないことが原因であると考えられる。今後はアボイドノートの考慮を検討したい。以上の結果より、経験者の演奏に対して僅かに音楽理論による補正を行うことは、ユーザの希望通りのメロディを推定することに有効であるが、過大な補正はユーザの意思とは異なるメロディを出力することに繋がる可能性があることがわかった。しかし、どの指定信頼度であっても、音楽理論をもとにした補正結果は、おおむね音楽的に違和感のないメロディを出力できることがわかった。

表 4.1: 指定信頼度のパターン (%)

パターン	$W_{pos-user}(\%)$	$W_{the-user}(\%)$
1	100	0
2	70	30
3	50	50
4	30	70
5	0	100

表 4.2: 音楽経験者による演奏に対する評価 (点)

音楽経験	A(1)	A(2)	B(1)	B(2)	C(1)	C(2)
ある (33 人)	2	3.24	3.60	3.63	2.64	1.91
ない (21 人)	3.14	3.62	3.05	2.91	3.00	2.01
平均 (54 人)	2.44	3.39	3.39	3.35	2.78	1.98

4.4 音楽未経験者による演奏

提案手法によるシステムが、ユーザのおおよそランダムな運指に対して、音楽理論により妥当なメロディを生成できるかどうか評価する。前節と同様に、楽器経験者による演奏と同じ条件で、楽器未経験者にも同様に4小節の曲を3曲演奏してもらい、図4.6に示すのD,E,Fの3種類の運指データを取得した。実際の運指データは、パターン1に相当する。演奏後に3種類の運指データに対し、表4.1の5パターンの指定信頼度を用いて必要に応じて補正を行う。なお、楽器未経験者に対して用いる指定信頼度として想定しているパターンは3, 4, 5の3種類であるが、本実験では音楽経験者用の指定信頼度を想定したパターン1, 2も適用して評価を行う。音楽経験者33人、音楽未経験者21人に音楽らしさを評価してもらった(表4.3)。

音楽理論の知識がない楽器未経験者による演奏では、ランダムな運指が行われたため、音楽理論によって補正される音が多く、補正パターンも楽器経験者による演奏時に比べて多かった。演奏者に演奏後の感想を聞いたところ、実際に曲をイメージして演奏を行ったわけではなく、思いつくがまま打鍵を行ったと話していた。そのため、指定信頼度がパターン1の時点では音列から調やコードを推定することが困難な運指も存在した。運指Dによる演奏では音楽経験者用の指定信頼度であるパターン1, 2による楽曲の評価が一番低かった(表4.3 D(1))。なお、音楽未経験者用

Figure 4.5 displays three musical patterns (A, B, and C) with two staves each. Red boxes highlight specific notes in the melody and chord tones.

- Pattern A:**
 - (1) Chords: FM7, Gm7, C7, FM7. The note G4 in the C7 chord is highlighted.
 - (2) Chords: FM7, Gm7, C7, FM7. The note G4 in the C7 chord is highlighted.
- Pattern B:**
 - (1) Chords: Fm7, B \flat 7, Fm7, B \flat 7. The note B \flat 4 in the B \flat 7 chord is highlighted.
 - (2) Chords: Fm7, B \flat 7, E \flat M7, Fm7. The note B \flat 4 in the B \flat 7 chord is highlighted.
- Pattern C:**
 - (1) Chords: FM7, Gm7, Gm7, C7. The note G4 in the first Gm7 chord and the note C5 in the C7 chord are highlighted.
 - (2) Chords: Dm7, G7, G7, CM7. The note D4 in the first Dm7 chord and the note C5 in the CM7 chord are highlighted.

図 4.5: 音楽経験者による運指と補正

の指定信頼度であるパターン 3, 4, 5 のうち、パターン 3, 4 においては、音楽経験者、音楽未経験者どちらからの評価も同様に高かったが、パターン 5 の時のみ評価が少し下がった。原因を検証するために、出力されたメロディとコードトーン、またはアボイドノートとの関係を調べた (図 4.7)。赤丸の箇所がコードトーンと一致している箇所であり、青四角の箇所が前節でも取り上げたアボイドノートの箇所である。パターン 5 によって補正された D(4) は、D(2), D(3) に比べてメロディとコードトーンが一致している数が少なく、またアボイドノートを用いている数が多いことがわかった。D(1) と D(4) ではメロディとコードトーンが一致している数とアボイドノートの数が等しいが、D(1) において推定された調は一般的によく使われる調ではなかったことが低評価の原因の一つであると考えられる。この結果より、音楽未経験者を想定して打鍵位置の重みを抑えることで、おおよそランダムな運指に対し

でも音楽的に優れたメロディを出力可能であることがわかった。しかし、アボイドノートを考慮した音選択なしには音楽性を損なう可能性があることもわかった。運指Eによる演奏では、音楽経験者と音楽未経験者のアンケート結果に差があった。音楽経験者はパターン1, 2, 3によって補正された楽曲を高く評価し、音楽未経験者は全てのパターンにおいて大きな差はなかった。これも運指Dの時と同様にメロディとコードトーンまたはアボイドノートとの関係を調べたところ(図4.7), E(1), E(2)に比べてE(3)で出力されたメロディにコードトーンが少ないことがわかった。また, E(1), E(2)ではGの音がD7に対するアボイドノートに当たるが, 次の小節のFM7に音が被っている。次の小節のFM7にとってGの音を9thの音と考えれば, この音はジャズの演奏時にもよく用いられる音であるため, Gの音は違和感を与える音にはならない。この結果より, 運指Dと同様にメロディとコードトーン的一致率は楽曲の音楽性に関わる重要な要素であることが考えられる。よって, 音楽理論による補正を行う場合, コードトーンの音に補正を行うことで音楽性を損なわない楽曲を出力できる可能性が高いことがわかった。運指Fによる演奏では, 音楽経験者と音楽未経験者のアンケート結果が一致して, パターン5による補正が最も高く評価された。前述の通り本研究で用いる手法では途中で調が変更されることはないものとしてコードの推定を行っている。パターン1, 2, 3, 4によって補正を行っている途中の2小節目までは推定調はFメジャーであったが, 3小節目の音から黒鍵の音が多くなった結果, Fメジャーであるという推定が破綻し, 最終的に無理のあるコードが選択されてしまった。このコード進行はどのダイアトニックコードにも当てはまっておらず, 音楽的に不安定な印象を与えてしまう。パターン5による補正では最初の推定調であるFメジャーに当てはまるように, かつコードトーンに合うメロディラインに補正を行っているため, 音楽的に違和感の少ない楽曲が出力された。この結果により, ダイアトニックスケールに沿ったコード付けを行い, 更にそのコードトーンと一致するメロディになるよう補正を行うことで, 音楽性を損なわない楽曲を出力できることがわかった。以上の結果により, 音楽未経験者によるランダムな運指を音楽的に優れた楽曲に補正するためには, ダイアトニックコードに沿ったコード付けを行い, 更にそのコードトーンと一致するメロディになるよう補正を行うことで, 音楽性を損なわない楽曲を出力できる可能性が高いことがわかつ

た。なお、メロディの補正を行う場合には、その音がダイアトニックスケール上の音であったとしてもアボイドノートである可能性があるため、アボイドノートになる確率を下げなければならないということがわかった。また、アンケートでは音楽経験の全くない人がランダムな運指で演奏したメロディは、コードの進行感が薄く、伴奏に違和感を感じやすくなるような気がした、という感想もあった。ダイアトニックコードに沿ったコード付けを行うことでこの問題を解決することが期待できる。

The figure displays musical notation for three keys: D, E, and F. Each key section contains four patterns (パターン) of chords and notes. Red boxes highlight specific notes in the patterns.

D Key:

- パターン 1, 2: Chords D7, GM7, Am7, D7. Notes: G, A, B, C, D, E, F, G.
- パターン 3: Chords C7, FM7, C7, FM7. Notes: C, D, E, F, G, A, B, C.
- パターン 4: Chords C7, FM7, C7, FM7. Notes: C, D, E, F, G, A, B, C.
- パターン 5: Chords C7, FM7, Gm7, C7. Notes: C, D, E, F, G, A, B, C.

E Key:

- パターン 1, 2: Chords D7, FM7, GM7, Am7. Notes: D, E, F, G, A, B, C, D.
- パターン 3: Chords D7, FM7, CM7, Dm7. Notes: D, E, F, G, A, B, C, D.
- パターン 4, 5: Chords B^bM7, B^bm7, E^b7, Dm7. Notes: B, C, D, E, F, G, A, B.

F Key:

- パターン 1, 2, 3: Chords G^b7, D^bm7, BM7, Dm7. Notes: F, G, A, B, C, D, E, F.
- パターン 4: Chords G^bM7, EM7, BM7, EM7. Notes: F, G, A, B, C, D, E, F.
- パターン 5: Chords C7, Gm7, FM7, Gm7. Notes: F, G, A, B, C, D, E, F.

図 4.6: 音楽未経験者による運指と補正

表 4.3: 音楽未経験者による演奏に対する評価 (点)

音楽経験	D(1)	D(2)	D(3)	D(4)	E(1)	E(2)	E(3)	F(1)	F(2)	F(3)
ある (人)	1.88	3.18	3.30	2.85	2.76	2.73	2.39	1.39	1.79	2.79
ない (人)	2.48	3.19	3.33	2.90	2.95	3.05	3.05	2.14	2.38	3.81
平均 (人)	2.11	3.19	3.31	2.87	2.83	2.85	2.65	1.69	2.02	3.19

Figure 4.7 shows musical notation for three patterns (D, E, F) across three variations (1, 2, 3). Red circles highlight specific notes, and blue boxes highlight specific chords. Chords are labeled above the notes.

Pattern D:

- (1) Chords: D7, GM7, Am7, D7. Notes: G4, A4, B4, C5.
- (2) Chords: C7, FM7, C7, FM7. Notes: G4, A4, B4, C5.
- (3) Chords: C7, FM7, C7, FM7. Notes: G4, A4, B4, C5.
- (4) Chords: C7, FM7, Gm7, C7. Notes: G4, A4, B4, C5.

Pattern E:

- (1) Chords: D7, FM7, GM7, Am7. Notes: G4, A4, B4, C5.
- (2) Chords: D7, FM7, CM7, Dm7. Notes: G4, A4, B4, C5.
- (3) Chords: B^bM7, B^bm7, E^b7, Dm7. Notes: G4, A4, B4, C5.

Pattern F:

- (1) Chords: G^b7, D^bm7, BM7, Dm7. Notes: G4, A4, B4, C5.
- (2) Chords: G^bM7, EM7, BM7, EM7. Notes: G4, A4, B4, C5.
- (3) Chords: C7, Gm7, FM7, Gm7. Notes: G4, A4, B4, C5.

図 4.7: 音楽未経験者による運指によるメロディとコードトーンまたはアボイドノートとの関係性

第5章 むすび

音楽経験者にとって最も自然な演奏方法は、鍵盤を見ずにタッチタイプで行うことである。タッチタイプでピアノを演奏できる理由として、ピアノの鍵盤から得られる垂直方向からの力覚フィードバックと、隣の鍵盤から得られる水平方向からの力覚フィードバックによって、自身の指の位置を判断し、次の鍵盤を打鍵するときは無意識的に指のずれを補正していることが考えられる。しかし既存のVRピアノは、ディスプレイなどの視覚情報を頼りに鍵盤の位置を把握して演奏を行うため、音楽経験者が普段から行う自然な演奏方法とは異なる。一方で、作曲支援システムでは自然な旋律、すなわち音の並びを作るため、音楽理論を用いている。そこで本研究では、従来のVRピアノに対し、作曲支援システムでも用いられている音楽理論を利用して誤打鍵による不自然な音の並びを補正することで、タッチタイプ演奏が可能なVRピアノシステムを提案した。また、提案システムでは、ユーザの打鍵位置の信頼度合い、あるいは音楽理論の重視度合いの重み付けを調整できる。音楽理論の信頼度を高めることにより、音楽未経験者がランダムな運指で演奏を行っても適切なメロディを生成することが可能である。評価実験では音楽経験者と音楽未経験者それぞれにシステムを用いて演奏してもらい、信頼度の比を変更しつつ補正を行った。補正結果のメロディを評価してもらうことで本システムの有用性を示したと同時に現在の問題点を発見することができた。

今後の課題として、アボイドノートの考慮や、ダイアトニックコードに忠実なコード選択とそのコードトーンから外れない補正方法の検討などが挙げられる。また、現段階では演奏を行うときにテンポや拍子を固定して行っているため、演奏者はシステムから提示されたテンポと拍子以外で曲を演奏することができない。今後は運指からテンポと拍子を推定する手法も検討したい。他にも現段階では考慮していなかったマイナーキーとブルースの曲の考慮や、Yuriら [16] によって示された Leap

Motion の問題点を解消するため、カメラベースでの運指の取得を行うシステムの構築や、データベースから数学的に導き出された調やコード進行にランダム性を持たせることが挙げられる。現在のシステムでは、打鍵の位置や順番によって調やコードを決定しており、同じ運指で演奏を行った場合、全く同じ補正が行われる。なお、この時に選択される調やコードはサンプルデータ内で多く用いられる調やコードが選択されるため、様々な運指で演奏を行ったとしても生成した楽曲が似通ったものになりがちである。そこで、3章にて説明した調確率、コード確率、鍵盤確率が最も高いものを推定調、推定コード、打鍵音として決定するのではなく、これらの確率をそのままその調、コード、鍵盤が選択される可能性に変換し、そこからランダムで調、コード、鍵盤を選択する手法も検討したい。この手法を取り入れることによって、これまで選択されなかったが音楽理論的に優れていると判定される他の調やコードや打鍵音が選択される可能性が高くなり、より面白味のある楽曲を作成できるようになるのではないかと考えている。

謝辞

本研究を進めるにあたり、ご指導を頂いた名古屋工業大学准教授 舟橋健司先生、助教 伊藤宏隆先生に心から感謝致します。また、日常の議論を通じて多くの知識や示唆を頂いた名古屋工業大学 舟橋研究室所属の学生の皆様に深く感謝し、厚く御礼申し上げます。本論文を執筆するにあたり、アンケートにご協力いただいた全ての皆様に厚く感謝申し上げます。

参考文献

- [1] Chia-Hung Yeh, Wen-Yu Tseng, Jia-Chi Bai, Ruey-Nan Yeh, Sun-Chen Wang, Po-Yi Sung, “Virtual Piano Design via Single-view Video based on Multifinger Actions Recognition”, *Human-Centric Computing*, ID: 11517146, 2010.
- [2] 竹川佳成, 寺田努, 塚本昌彦, “運指認識技術を活用したピアノ演奏支援システムの構築”, *情報処理学会論文誌*, No. 52, Vol. 2, pp.917-927, 2011.
- [3] Mikel Sagardia, Katharina Hertkorn, David Sierra Gonzalez, Claudio Castellini, “Ultrapiano: A Novel Human-Machine Interface Applied to Virtual Reality”, *International computer music association*, Vol. 2014, 2014.
- [4] Claudio Castellini, Katharina Hertkorn, Mikel Sagardia, David Sierra Gonzalez, Markus Nowak, “A virtual piano-playing environment for rehabilitation based upon ultrasound imaging”, *IEEE International Conference on Biomedical Robotics and Biomechatronics*, ID: 15143949, 2014.
- [5] Enkhtogtokh Togootogtokh, Timothy K. Shih, W.G.C.W.Kumara, Shih-Jung Wu, Shih-Wei Sun, Hon-Hang Chang, “3D finger tracking and recognition image processing for real-time music playing with depth sensors”, *Multimed Tools Appl 2017*, ID: 1573-7721, 2017.
- [6] Apple Inc., GarageBand, <https://www.apple.com/jp/ios/garageband/>, 2017
- [7] Wegrass Interactive, Touch Piano, <http://grizzly.wegrass.com/?type=app-dock>, 2017
- [8] ジャン・ル・ロン・ダランベール, 片山千佳子, 安川智子, 関本菜穂子, “ラモー氏の原理に基づく音楽理論と実践の基礎”, 春秋社, 2012

- [9] 川島奨大, 柳英克, “初学者を対象とする作曲支援 Web アプリケーションシステムの開発と評価”, *IPSJ Interaction*, ID: 1-405-54, 2017
- [10] 岩田拓也, 澤野弘明, 池田輝政, 菱田隆彰, “カードを用いた作詞作曲支援システムの提案”, 情報処理学会第 75 回全国大会, 1ZG-7, 2013
- [11] Chika Matsusue, Kenji Funahashi, Shinji Mizuno, “Touch-typable VR piano that corrects positional deviation of fingering based on music theory”, *SIGGRAPH '17 Posters Article No.33*, 2017
- [12] Adam Goodwin, Richard Green, “Key Detection for a Virtual Piano Teacher”, *IVCNZ*, ID: 14062333, 2014.
- [13] Jernej Barbic, Doug James, “Six-dof haptic rendering of contact between geometrically complex reduced deformable models”, *IEEE Transactions on Haptics*, No. 1, Vol. 1, pp.39-52, 2008.
- [14] Bullet Physics. Bullet, <http://bulletphysics.org/>, 2017.
- [15] Inwook Hwang, Hyungki Son, Jin Ryong Kim, “AirPiano: Enhancing Music Playing Experience in Virtual Reality with Mid-Air Haptic Feedback”, *IEEE World Haptics Conference*, ID: 17061639, 2017.
- [16] Yuri De Pra Fausto Spoto, Federico Fontana, Linmi Tao, “Infrared vs. Ultrasonic Finger Detection on a Virtual Piano Keyboard”, *IEEE World Haptics Conference*, pp. 654-658, 2014.
- [17] Takayuki Hoshi, Masafumi Takahashi, Takayuki Iwamoto, Hiroyuki Shinoda, “Noncontact tactile display based on radiation pressure of airborne ultrasound,” *IEEE World Haptics Conference* No. 3, Vol. 3, pp. 155-165, Jul. 2010.
- [18] Tom Carter, Sue Ann Seah, Benjamin Long, Bruce Drinkwater, Sriram Subramanian, “Ultrahaptics: multi-point mid-air haptic feedback for touch surfaces.” *UIST*, S. Izadi, A. J. Quigley, I. Poupyrev, and T. Igarashi Eds. ACM, pp. 505-514, 2013,

- [19] Takayuki Yoshida, Teruhisa Hochin, Hiroki Nomiya, “Automatic Composition Using Melodic Outlines and Music Theory”, *SNPD 2017*, pp. 525-530, 2017
- [20] Kouji Ishizaka, Chord progression stories that can be sung soon, Rittor Music Inc., 2009
- [21] Tetsuro Kitahara, Yuichi Tsuchiya, “Short-term and Long-term Evaluations of Melody Editing Method based on Melodic Outline”, *International Computer Music Conference (ICMC)* , *Sound and Music Computing Conference (SMC)*, pp.1204-1211, 2014
- [22] 池田輝政, 菱田隆彰, “劇伴の作成を支援する楽曲作成システム「Lazy Composer」”, 情処研報, 音楽情報科学, 2011-MUS-91. No. 10, pp. 1-4, 2011
- [23] Andrew Milne, The Tonal Centre, <http://www.tonalcentre.org/>, 2017
- [24] 納浩一, “ジャズ・スタンダード・バイブル ～セッションに役立つ不朽の227曲～”, Rittor Music Inc., 2010
- [25] Leap Motion, Leap Motion, <https://www.leapmotion.com/>, 2016
- [26] Leap Motion, Leap Motion API ドキュメント, <https://developer.leapmotion.com/documentation/cpp/>, 2016
- [27] Technimo LLC, iReal Pro, <https://itunes.apple.com/jp/app/ireal-pro/id298206806?mt=8>, 2018

付録A 本論文で用いられる音楽用語

本章で紹介する音楽用語，音楽理論は The Tonal Centre[23] とラモー氏の原理に基づく音楽理論と実践の基礎 [8] を参考にしている。

A.1 音名，階名

- 音名：絶対的な音の高さを表す。日本では音名にハニホヘトイロハを用いる。他にも英語，ドイツ語などが用いられる。
 - 日本音名：ハ，ニ，ホ，ヘ，ト，イ，ロ
 - アメリカ音名：C(シー), D(ディー), E(イー), F(エフ), G(ジー), A(エー), B(ビー)
 - ドイツ音名：C(ツェー), D(デー), E(エー), F(エフ), G(ゲー), A(アー), H(ハー)
- 階名：相対的な音の高さを表す。一般的にはイタリア語のドレミファソラシドを用いる。この場合，基準となる音は常にドである。

A.2 調，音程

- 音程：2つの音の隔たりや間隔の大きさを音程と呼び，音程の大きさにより呼称が存在する (表 A.1)。基準となる音から1オクターブ上の基準となる音の間には12個の半音が含まれる。また，全音とは半音2個分の音程のことである。
- ♯(シャープ)：基準となる音から半音分の音程を上げる記号。
- ♭(フラット)：基準となる音から半音分の音程を下げる記号。

表 A.1: 音程

音程	表記	概要
完全1度(1)	I	基準となる音と同じ音程
短2度(-2)	II \flat	1つの半音からなる音程
長2度(2)	II	1つの全音からなる音程
短3度(-3)	III \flat	1つの半音と1つの全音からなる音程
長3度(3)	III	2つの全音からなる音程
完全4度(P4)	IV	1つの半音と2つの全音からなる音程
増4度, 減5度(+4,-5)	V \flat	3つの全音からなる音程
完全5度(P5)	V	1つの半音と3つの全音からなる音程
増5度, 短6度(+5,-6)	VI \flat	4つの全音からなる音程
長6度(6)	VI	1つの半音と4つの全音からなる音程
短7度(-7)	VII \flat	5つの全音からなる音程
長7度(7)	VII	1つの半音と5つの全音からなる音程
完全8度(P8)	I	6つの全音からなる音程, 基準となる音の1オクターブ上の音程

- 調, キー: メロディや和音がある基準音と調性を持ちつつ構成される場合, この音組織を調と呼ぶ. 例えば基準音がCである場合, C,D,E,F,G,A,Bによって構成される調を, Cメジャーキーと呼ぶ. また, C,D,E \flat ,F,G,A \flat ,B \flat によって構成される調をCマイナーキーと呼ぶ.
 - メジャー: 音程が全音, 全音, 半音, 全音, 全音, 全音, 半音の音で構成された調.
 - マイナー: 音程が全音, 半音, 全音, 全音, 半音, 全音, 全音の音で構成された調.

A.3 コード

- 和音, コード: 同時に聞こえる3つ以上の高さの異なる音の混合. 本論文で主に使われるコードは以下の通りである(表A.2). また, コードの構成音のことをコードトーンと呼ぶ.
- 和声, ハーモニー: 複数の和音の連なり.

表 A.2: コード

コード	Cを基準音とした場合の構成音と度数
CM7	C(1), E(3), G(P5), B(7)
Cm7	C(1), Eb(-3), G(P5), Bb(-7)
C7	C(1), E(3), G(P5), Bb(-7)
CmM7	C(1), Eb(-3), G(P5), B(7)
Csus4	C(1), F(P4), G(P5)
Cdim	C(1), Eb(-3), Gb(-5), A(6)
Cm7-5	C(1), Eb(-3), Gb(-5), Bb(-7)

- **ダイアトニックコード**: ダイアトニックスケール (付録 A.4) 上の音で構成されたコード。例えばCメジャーキーにおけるダイアトニックコードはCM7, Dm7, Em7, FM7, G7, Am7, Bm7-5である。

A.4 スケール

- **音階, スケール**: 高さ順に並べられたオクターブ以内の階段状の音列。なおスケールの音は以下の3種類に分類される。
 - **コードトーン**: 付録 A.3
 - **テンションノート**: コードの響きに緊張感を与え、かつコード進行を阻害しない音。
 - **アボイドノート**: コードトーンと同時に鳴らした時に音がぶつかり、コードのサウンドや機能を崩してしまう音。
- **調号**: 各調のスケールの構成音を示す。フラットやシャープの数で表現する。例えば調がFメジャーの調号はフラットが1つである。
- **ダイアトニックスケール**: ある調の音階上の7音で構成されたスケール。例えばCメジャーキーにおけるダイアトニックスケールはC, D, E, F, G, A, Bである。また、メジャーキーのダイアトニックスケール上の音のスタート位置によって、7種類のスケールが存在する (表 A.3)。

表 A.3: スケール

スケール名	C メジャースケールにおける構成音と基準音からの度数
Ionian	C(1), D(2), E(+3), F(P4), G(P5), A(6), B(+7)
Dorian	D(1), E(2), F(-3), G(P4), A(P5), B(6), C(-7)
Phrygian	E(1), F(-2), G(-3), A(P4), B(P5), C(-6), D(-7)
Lydian	F(1), G(2), A(+3), B(+4), C(P5), D(6), E(+7)
Mixolydian	G(1), A(2), B(+3), C(P4), D(P5), E(6), F(-7)
Aeolian	A(1), B(2), C(-3), D(P4), E(P5), F(-6), G(-7)
Locrian	B(1), C(-2), D(-3), E(P4), F(-5), G(-6), A(-7)

発表論文リスト

投稿論文

1. Chika Matsusue, Kenji Funahashi, Shinji Mizuno, "Touch-typable VR piano that corrects positional deviation of fingering based on music theory", SIGGRAPH '17 Posters Article No.33, 2017