

平成 29 年度 卒業論文

カラオケのための
ハモリパート練習システム
～ハモリパートの自動生成および
練習システムの試作～

指導教員 北原 鉄朗 准教授

日本大学文理学部情報科学科

白石 美南

2018 年 2 月 提出

概 要

カラオケの楽しみ方の1つに、主旋律と副旋律を2人で歌ってハーモニーを奏でることがあげられる。しかし、副旋律（いわゆるハモリパート）を習得するには多くの問題が発生する。例えば、副旋律を練習するためのお手本を探すのに手間がかかってしまうこと、主旋律歌唱者の特性（キー）を考慮せず練習を行った場合、通常原曲の副旋律を練習するだけでは適切なハーモニーを奏でられないこと、ユーザが個人で練習を行った場合、実際にハーモニーを奏でる際につられてしまうことといった問題がある。このことから本研究は、ユーザが副旋律を練習する際に生じる、“副旋律を担当するユーザが副旋律を入手することが困難”、“主旋律歌唱者の特性（キー）を考慮できない”、“主旋律につられてしまう”の3つの問題の解消を狙ったものである。

まず、副旋律を担当するユーザが副旋律を入手することが困難であるという問題については副旋律の自動生成を行い、副旋律歌唱者が手本となる副旋律を探す・生成する手間を無くすことで解決する。なお、副旋律の自動生成は隠れマルコフモデル（HMM）、ルールベースの2手法で行う。HMMでは、主旋律における各音名を状態とし、主旋律および伴奏を観測系列としたHMMで、楽曲の調をもとにした出現確率、主旋律の音名の遷移確率、主旋律をもとにした出力確率、伴奏をもとにした出力確率を利用している。ルールベースでは、副旋律は主旋律と3、4度の関係を取る場合が多いことから、楽曲の伴奏を考慮して主旋律と3、4度の関係を副旋律として配置する。

次に、生成した副旋律を利用した副旋律練習システムの実装を行う。このとき、主旋律歌唱者の特性（キー）を考慮できない問題、主旋律につられてしまう問題を解決するために、2つの機能を実装した。1つ目の機能としては、主旋律歌唱者の特性（キー）を考慮できない問題を解決するために、練

習楽曲の主旋律，副旋律，伴奏の音高を変更できる機能を実装した．これによって，相手の歌唱者のキーに合った音高で副旋律を練習することができる．2つ目の機能としては，主旋律につられてしまう問題を解決するために，自身の歌唱した音高を視覚化した状態でフィードバックする機能を実装した．これによって，ユーザは視覚化された自分の音高を，適切な音高に調節しながら歌唱することができる．

実装した副旋律練習システムの有効性を確認するために，まずは副旋律歌唱者が手本とする副旋律が適切なものであるかどうか検証する．音楽大学の大学院生が作成した正解データとの一致率を求めたところ，HMMで生成した副旋律については，上八モリが平均52.3%，下八モリが平均53.6%となり，ルールベースで生成した副旋律については，上八モリが平均55.5%，下八モリが平均64.2%となった．これらの一致率は正解データとの比較であるので，一致しなかった場合でも音楽的に不適切であるとは限らない．このことから，各手法において特に一致率の低い11曲（HMM：6曲，ルールベース：5曲）に対して音楽的な適切さの評価を別の音楽の専門家に依頼した．その結果，11曲全曲において67%を超える評価となったため，各手法で生成した副旋律は一定の音楽的妥当性があることが分かった．

目 次

目 次	iii
図目次	v
表目次	vii
第1章 序 論	1
1.1 背景	1
1.2 目的	2
1.3 論文の構成	3
第2章 関連研究	5
2.1 自動和声付け (副旋律自動生成)	5
2.1.1 芸大和声 [3]	5
2.1.2 ルールベースによる合唱譜自動生成システム [4]	5
2.1.3 ルールベースによるポップス音楽向けの自動和声付け [5]	6
2.1.4 ベイジアンネットワークを用いた四声体和声付け [6]	6
2.1.5 隠れマルコフモデルを利用した自動和声付け [7]	6
2.2 歌唱練習	7
2.2.1 相対音感習得システム [8]	7
2.2.2 歌唱力向上支援インタフェース [9]	7
2.3 研究課題	8
2.3.1 自動和声付け	8
2.3.2 副旋律練習	9

2.3.3	研究課題設定	9
第 3 章	副旋律自動生成	11
3.1	HMM による手法	11
3.2	自動副旋律付け：ルールベースによる手法	13
第 4 章	副旋律自動生成手法の評価	15
4.1	評価方法	15
4.2	評価結果	15
4.3	考察	18
4.3.1	副旋律評価	18
4.3.2	生成された副旋律の主観評価	20
第 5 章	副旋律歌唱練習システムへの応用	23
5.1	システムの流れ	23
5.2	相手の特性 (キー) に合わせた練習	25
5.3	主旋律につられない練習	25
5.4	その他の機能	26
第 6 章	結 論	29
	参考文献	31

目 次

4.1	HMM で生成した上八モリの 85 曲全体の精度	16
4.2	ルールベースで生成した上八モリの 85 曲全体の精度	16
4.3	HMM で生成した下八モリの 85 曲全体の精度	17
4.4	ルールベースで生成した下八モリの 85 曲全体の精度	17
4.5	85 曲中で同じ音名が連続する箇所の出現回数 (「上で八モる」 場合)	19
4.6	85 曲中で同じ音名が連続する箇所の出現回数 (「下で八モる」 場合)	19
5.1	曲選択画面	24
5.2	練習画面	24
5.3	ユーザの音高表示	28

表 目 次

4.1	副旋律生成結果の精度	16
4.2	再評価をする HMM で生成した副旋律の精度	21
4.3	再評価するルールベースで生成した副旋律の精度	21
4.4	HMM で生成した副旋律の再評価結果	22
4.5	ルールベースで生成した副旋律の再評価結果	22

第1章 序 論

1.1 背景

コンピュータやインターネット技術の発達により、音楽の配信や配布が容易になり、我々の身近なものとして浸透するようになった。また、音楽の楽しみ方についてもカラオケやミュージカル鑑賞といったように、自ら歌唱して楽しむ方法や鑑賞する方法もある。音楽の楽しみ方は人によって様々であるが、特にカラオケは安価で手軽に通えるため、多くの人々が利用している。カラオケをより楽しむための研究も数多く存在する。栗原らの研究 [1] では、カラオケでの歌い手と聞き手の双方間のコミュニケーションをより深めるためのタンバリンを用いた支援を行う提案をしている。楽曲に合わせたタンバリン譜を自動で生成し、タンバリンの演奏によって聞き手がカラオケに能動的に参加できる環境を作っている。これにより、歌にあわせて聞き手がタンバリンを演奏することができ、歌い手と聞き手の双方が楽しむことができる。浦川らの研究 [2] では、歌が上手に歌えない音痴の人たちが付き合いによりカラオケで歌唱することになった時に、歌唱時の音高のずれを修正してくれるシステムを提案している。ただし音高のずれを音高変換のみで修正すると、歌声が機械的になる問題があるため、音高変換後、「歌声らしさ」を与えるピッチ変動を行って音声合成を行い、より自然に音高の修正を行うという研究である。

カラオケは、通常は、ユーザが専用の機器から流れる伴奏に合わせて主旋律を歌唱し、楽しむものである。最近では、カラオケの楽しみ方として主旋律を担当するユーザと、副旋律を担当するユーザに分かれ、歌唱者同士でハーモニーを奏する方法もある。ハーモニーを奏するためには主旋律、副旋律とも

に正確に歌唱することが要求され、カラオケで歌唱する以前に練習を多く行う必要がある。しかし、副旋律を練習するための環境を整えるためには手間がかかるといった問題がある。主旋律を担当するユーザが練習をする場合、インターネットで歌唱する曲を検索するだけで、担当部分の参考になる動画などを容易に探し出すことができる。しかし、副旋律のみを歌唱した動画などがインターネット上にあがることは稀であり、担当部分の参考になるものを探すことが困難であるという問題がある。そのため、副旋律を担当するユーザの多くは、探すことが容易な主旋律と副旋律を合わせて歌唱したものから、副旋律のみを聞き分けることで練習を行っている。しかし、音楽的知識のないユーザには主旋律と副旋律を聞き分けることが困難である。また、原曲に副旋律の存在しない曲が存在する。このような曲でハーモニーを奏でる場合、ユーザ自身が適切な副旋律を自ら考案し、それを練習する方法があるが、この方法も音楽的知識のないユーザには困難である。

ユーザが参考にすることができる副旋律を発見・生成できた場合でも、以下の問題が発生する。実際にユーザ同士でハーモニーを奏でる際には、主旋律歌唱者が原曲キーで歌うとは限らない。そのため、主旋律歌唱者の特性（キー）に合わせた練習を行わなければならない問題がある。主旋律を歌唱しているユーザにつられてしまうという問題も存在する。

1.2 目的

1.1 節で述べたように副旋律を練習するためには以下の3つの問題が存在する。

- 副旋律を担当するユーザが副旋律を入手することが困難であるという問題
- 主旋律歌唱者の特性（キー）を考慮できない問題
- 主旋律につられてしまう問題

本研究は、これら 3 つの問題を解決するハモリ練習システムを実現することを目的とする。具体的にはこのシステム内で以下の 3 つを実現する事を目的とする。

- 自ら副旋律を発見できない人のための副旋律自動生成
- 主旋律歌唱者の特性（キー）に合わせた練習を行うための機能の実装
- 主旋律を歌唱しているユーザにつられないようにする練習を行うための機能の実装

本研究では、特に 1 つ目の目的に焦点を当てて研究を進める。

1.3 論文の構成

本論文の構成は次の通りである。2 章では自動和声付けに関する研究事例、歌唱練習システムに関する研究事例についてそれぞれ述べ、本研究と関連研究との差分を挙げた上で、設定した研究課題について述べる。3 章では 2 章で設定した研究課題の 1 つ目を達成するための副旋律自動生成手法について述べる。4 章では本研究で提案した手法によって生成された副旋律の評価をするために行った実験について述べる。5 章では 2 章で設定した研究課題の 2 つ目、3 つ目を達成するために、生成した副旋律を利用し実装した副旋律歌唱練習システムについて述べる。6 章では本論文の結論を述べる。

第2章 関連研究

1.2 節で述べた 3 つの目的を達成するために、まず既存の副旋律を自動生成する手法および歌唱練習システムに関する研究について述べる。

2.1 自動和声付け (副旋律自動生成)

本節では、副旋律の自動生成手法について述べられている研究事例を紹介する。

2.1.1 芸大和声 [3]

芸大和声 [3] とは多くの日本の作曲家が、和声付けをするための基礎教育を行う際に利用している書籍である。長調、短調に基づく調和和声音楽である西洋古典音楽の和声理論や、楽曲における複数音を同時に発する際の、音同士の協和・不協和、連結に関する規則の体系化を行う手法について述べられている。和声付けをルールベースで行う研究の多くは、この書籍を参考にしている。

2.1.2 ルールベースによる合唱譜自動生成システム [4]

エバンズらの研究 [4] では、芸大和声 [3] で述べられているバス課題の 3 規則を適用した、合唱譜自動生成システムの実装を行っている。バス課題とは、与えられたバス譜の上にソプラノ・アルト・テノール (上 3 声) を和声の諸規則に反せずに配置する課題である。芸大和声ではバス課題のルールとして「禁ぜられる和声配置」、「推奨される和音配置」、「標準的な連結の原則」の

3 規則を挙げている。エバンズらの研究は、バス譜から上記の3種類の規則を満たす和声(上3声)を配置するシステムである。

2.1.3 ルールベースによるポップス音楽向けの自動和声付け [5]

三浦らの研究 [5] は、ポピュラー音楽を対象にユーザの自作旋律に対する和声および伴奏の付与を行っている。ポピュラー音楽系の旋律に対応するために、対応する和音を古典西洋音楽の規範 [3] からポップス用に拡張を行っている。また、旋律に対しては刺繍音、経過音、先行音などのアプローチノート処理を施すようにルールを定めている。

2.1.4 ベイジアンネットワークを用いた四声体和声付け [6]

鈴木らの研究 [6] では、ソプラノ譜を入力として、アルト・テノール・バス(下3声)を配置するシステムの実装を行っている。四声体和声(ソプラノ・アルト・テノール・バスによる4つの声体)の和声付けを行う際に、ベイジアンネットワークを用いることで、音楽的同時性と音楽的連続性の両立を実現している。音楽的同時性とは、声部間の縦のつながりのことで不協和音を作らないことである。音楽的連続性とは、声部内の横のつながりのことで和声の流れを自然にすることである。

2.1.5 隠れマルコフモデルを利用した自動和声付け [7]

菅原らの研究 [7] は、音声認識と自動和声付けを同型の問題として考え、音声認識で使用されている確率モデルを利用した確率的な推定問題として自動和声付けを行っている。各音の音高の推移や小節内での拍位置に注目して旋律モデルにする。和音進行を和声常套句(短い和声の進行列)の n -gram モデル化を行う。この2つのモデルを隠れマルコフモデル(以下 HMM とする)として統合し、モデルのパラメータは既存の楽曲の統計から適切な値を定める

ことができる．和声付けは，与えられた旋律に対して HMM の状態遷移系列を求める事後確率最大化問題として定式化されている．

2.2 歌唱練習

本節では，歌唱練習システムの実装を行っている研究事例を紹介する．

2.2.1 相対音感習得システム [8]

福本らの研究 [8] は，ボーカルピッチマッチング能力保持者を対象とした，相対音感の習得を行うための歌唱練習システムである．相対音感とは，ある音を基準にして相対的に他の音を聞き分ける能力のことである．歌唱，楽器の演奏時には，「ハーモニー」が主体となるため，「絶対音感」と比較して「相対音感」が重要であるとされる．ユーザ本人の声を参照音源として使用し，自身の歌唱している音程を視覚的に見えるようにフィードバックする機能や，自分の発声した歌声を聴覚的にフィードバックする機能を利用して練習することで，ユーザは相対音感を身に付けることができる．システムから発生される単音が，自身の発生している音とどの程度音程差があるかどうかを把握する能力を上げる支援も行っている．

2.2.2 歌唱力向上支援インタフェース [9]

この研究では，ユーザが歌を歌ったり聴いたりするだけでなく，「見る」歌唱力向上支援インタフェースを提案している．中野らの研究 [9] は，ユーザーの歌唱を録音し，歌唱の音高を軌跡として可視化する．ビブラートなどは着色され，歌い方がおかしかった部分を視覚的に把握することができる．また，CD 音源と，自分の歌唱を聴き比べる機能も備えている．CD 音源のボーカルを，伴奏を含む楽曲から自動で抽出し，基本周波数 (F_0) を推定する．マウス操作で推定結果を調節することでボーカルの抽出精度を向上させる機能も持っている．

2.3 研究課題

1.2節で述べたように、ユーザが副旋律を練習するためには、副旋律を担当するユーザが副旋律を入手することが困難であるという問題、主旋律歌唱者の特性（キー）を考慮できない問題、主旋律につられてしまう問題がある。

2.3.1 自動和声付け

本研究では、カラオケでよく歌われているポップスの曲を対象に、副旋律を担当するユーザが副旋律を入手することが困難であるという問題を、副旋律を自動生成することによって解決する。自動副旋律付けの方法として、伴奏がない主旋律に複数の声部を付与して3声部や4声部の和声を生成するもの[4, 6]がある。しかし、[4]のような主旋律（バス譜）からテノール、アルト、ソプラノ譜を生成しているものや、[6]のような主旋律（ソプラノ譜）からバス、テノール、アルト譜を生成しているものは、合唱曲を対象に和声付けを行っている。そのため、生成した3つの副旋律は同時に歌唱されることが前提となり、4つの旋律のうち1つでも抜けて歌唱した場合、適切なハーモニーを奏でられない可能性がある。[7]は古典西洋音楽を対象とした自動副旋律生成であり、伴奏なし主旋律に対してコード進行や副旋律を付与を行っている。しかし、[5]でポップスの楽曲への和声付けでは古典音楽の規範を少し緩める必要があると述べられているため、[7]をポップスに適用した場合、適切な副旋律が生成されない可能性がある。[4, 5, 6, 7]は、共通して楽曲の伴奏情報を考慮していない。そのため、生成された副旋律が伴奏と不協和音を生み出す可能性がある。[8]では単音に対する和音でしか練習することができないため、連続した音（楽曲等）の副旋律の練習には向いていない。

以上のことから、副旋律を担当するユーザが副旋律を入手することが困難であるという問題を解決するために、ポップスに適用できる副旋律の自動生成を行う。

2.3.2 副旋律練習

ユーザが参考にすることができる副旋律を発見・生成できた場合でも、まだ副旋律の練習を行うにはいくつか問題がある。1つ目として、主旋律歌唱者の特性によっては、発見・生成した副旋律を利用できない場合があるという問題がある。副旋律の練習を行うにあたって、副旋律を発見した場合でも、実際に主旋律を歌唱する相手によっては歌唱するキーが異なる場合がある。そのため、実際にカラオケで歌唱する際に主旋律歌唱者がキーの変更を行うと、発見した副旋律と旋律が異なるため、適切な練習を行うことができない。

2.3.3 研究課題設定

以上のことから副旋律歌唱者が歌唱練習を行えるようなシステムを実装する上で以下のような研究課題を設定する。

- ポップスに適用できる伴奏を考慮した副旋律の自動生成を行うこと。
- ハモリパート練習時に、主旋律を歌唱する相手の特性に合わせて副旋律を練習できるようにすること。
- 主旋律につられない練習をできるようにすること。

第3章 副旋律自動生成

音楽理論で3和音を配置する規則に着目すると、長3和音は、根音、根音から長3度の音、さらに短3度の3つの音が配置されている。同様に短3和音についても、根音、根音から長3度の音、さらに短3度の3つの音が配置されている。このように、和音を配置する際には各音と音の距離が3度の関係であることが基本となる。ポップス楽曲の主旋律に配置される副旋律の規則についても、主旋律の音に対して3度の関係の音を、副旋律として配置していることが多い。実際に4章の副旋律評価に使用したJ-pop正解データ85曲に対して、主旋律と副旋律の関係を確認したところ、80.04%が3度の関係になっていることが分かった。しかし単にこの規則に従い、すべての主旋律の音に対して3度の音を副旋律として配置すると、配置した副旋律と伴奏が不協和音となってしまう場合がある。この問題を解決するために、本研究では伴奏情報を考慮した副旋律自動生成を提案する。自動副旋律生成を行うにあたって、本研究では、HMMとルールベースの2手法で行う。以降では副旋律を主旋律よりも高くする場合を「上でハモる」、副旋律を主旋律よりも低くする場合を「下でハモる」とする。

3.1 HMMによる手法

主旋律の音符列を n_1, \dots, n_m 、副旋律の音名列を h_1, \dots, h_m とする。 m は0以上11以下の整数を取るものとし、各々の値が音名(C, C \sharp , ..., B)に対応するものとする。主旋律の各音符 n_i の発音区間における伴奏の特徴を多次元ベクトル b_i で表す。また、当該楽曲の調を k とする。このとき算出すべ

きものは,

$$P(h_1, \dots, h_m \mid n_1, \dots, n_m, \mathbf{b}_1, \dots, \mathbf{b}_m, k) \quad (3.1)$$

が最大になる h_1, \dots, h_m である．副旋律のマルコフ性を仮定し，事前に与えられる主旋律および伴奏における依存関係を省略すれば，式 (3.1) は

$$\begin{aligned} & P(h_1)P(h_1|k)P(n_1|h_1)P(\mathbf{b}_1|h_1) \cdots \\ & P(h_m|h_{m-1})P(h_m|k)P(n_m|h_m)P(\mathbf{b}_m|h_m) \end{aligned} \quad (3.2)$$

と書くことができる．これは，副旋律における各音名 h_i を状態とし，主旋律 $\{n_i\}$ および伴奏 $\{\mathbf{b}_i\}$ を観測系列とした HMM に相当する．ただし，状態遷移確率は調 k に依存する．このようにモデル化すると，主旋律 $\{n_i\}$ と伴奏 $\{\mathbf{b}_i\}$ が与えられたときに副旋律 $\{h_i\}$ を付与する処理は，このモデルに対して Viterbi アルゴリズムを適用することで実現することができる．

以下，各確率の計算方法について述べる．

- $P(h_1)$ は初期確率であり，等確率とする．
- $P(h_i|k)$ は調 k が与えられたときの各音名の出現確率であり， h_i が調 k におけるダイアトニックノートで，かつ主音に対して一度，三度，五度の関係にあるときには 0.15，それ以外のダイアトニックノートの場合は 0.1，ダイアトニックノートでないときは 0.03 とする．
- $P(h_i|h_{i-1})$ は副旋律における音名の遷移確率である．これは，副旋律は主旋律と類似した特徴を持つべきだという観点から，当該楽曲の主旋律から音名の遷移確率を計算し，これを用いる．
- $P(n_i|h_i)$ は主旋律と副旋律の関係を表す．最も典型的な方法は主旋律

の増四度下に副旋律を付けることである．そこで次式により定義する．

$$P(n_i|h_i) = \begin{cases} 0.014285714 & |n_i - h_i| = 0 \text{ のとき} \\ 0.028571429 & |n_i - h_i| = 1 \text{ のとき} \\ 0.064285714 & |n_i - h_i| = 2 \text{ のとき} \\ 0.214285714 & |n_i - h_i| = 3 \text{ のとき} \\ 0.285714286 & |n_i - h_i| = 4 \text{ のとき} \\ 0.214285714 & |n_i - h_i| = 5 \text{ のとき} \\ 0.071428571 & |n_i - h_i| = 7 \text{ のとき} \\ 0 & |n_i - h_i| \geq 8 \text{ のとき} \end{cases}$$

- $P(b_i|h_i)$ は副旋律と伴奏の関係を表す．伴奏の音の中に，副旋律の音に対して短2度や長2度などの関係の音が含まれると，不協和音の原因となる．そこで，次のように定義する． b_i を12次元ベクトル $b_i = (b_{i,0}, \dots, b_{i,11})$ とする． $b_{i,j}$ は，主旋律の音符 n_i の発音時刻から消音時刻までの区間において，対応する音名が発音している時間の割合を表す．例えば， n_i が二分音符でその区間の伴奏においてCが二分音符で，EとGが四分音符で演奏されていたら， $b_i = (1.0, 0.0, 0.0, 0.0, 0.5, 0.0, 0.0, 0.5, 0.0, 0.0, 0.0, 0.0)$ となる．副旋律に対して短2度や長2度の伴奏音が含まれるのを防ぐには， $|h_i - j| \leq 2$ (ただし $h_i \neq j$) を満たす j に対して $b_{i,j}$ ができるだけ0に近づけばよい．そこで， $|h_i - j| \leq 2$ かつ $h_i \neq j$ のとき， $P(b_{i,j}|h_i)$ が正規分布 $N(0.0, \sigma_1^2)$ (σ_1 は実験的に定めた定数) に従うとする．それ以外の j に対しては $P(b_{i,j}|h_i)$ は一様分布に従うとする．ただし，実装上の都合により， σ_2 を十分に大きな定数として正規分布 $N(0.5, \sigma_2^2)$ で一様分布を近似している．なお， $b_{i,0}, \dots, b_{i,11}$ は互いに独立とする．

3.2 自動副旋律付け：ルールベースによる手法

本研究では，ルールに基づいて副旋律の自動生成を行う処理は下記の通りである．

1. 楽曲を半小節ごとに分割して伴奏パートの分析を行う。分析の結果、得られるベクトルの系列 b_1, \dots, b_T の内容は前節と同じであるが、主旋律のリズムと同期しておらず、半小節ごとに分割して処理を行っている点が異なっている。
2. 「下で八モる」際に、主旋律の各音符 n_i に対して次の処理を行う。
 - (a) 音符 n_i に時刻が対応する要素を $\{b_1, \dots, b_T\}$ から取り出す (b_t とする)。
 - (b) $b_t = (b_{t,0}, \dots, b_{t,11})$ を値が大きい順ソートし、 b_t から上位4つを取り出し、それらの音名を取得する (取得した4つの音名の集合を C_i とする)。
 - (c) C_i に主旋律より長三度下、すなわち $n_i - 4$ が含まれていれば、その音名 $h_i = n_i - 4$ を副旋律の音名とする (「下で八モる」の場合)。
 - (d) C_i に主旋律より短三度下、すなわち $n_i - 3$ が含まれていれば、その音名 $h_i = n_i - 3$ を副旋律の音名とする。
 - (e) C_i に主旋律より完全四度下、すなわち $n_i - 5$ が含まれていれば、その音名 $h_i = n_i - 5$ を副旋律の音名とする。
 - (f) C_i に主旋律より長三度下、短三度下、完全四度下のいずれの音名も含まれていない場合、直前の音名から主旋律と並行に移動するように副旋律の音名を定める。すなわち、 $h_i = h_{i-1} + (n_i - n_{i-1})$ とする。

「上で八モる」の場合も同様に定義する。

第4章 副旋律自動生成手法の評価

本研究は、副旋律を担当するユーザが歌唱練習を行えるシステムを開発することが目的である。システムには提案する手法により生成した副旋律を使用するため、生成した副旋律がどれだけ適切であったかどうかを検証する必要がある。本章では、行った実験の手順・考察について述べる。

4.1 評価方法

実験に使用する楽曲数は85曲で、本実験において、適切な副旋律とするデータ（以降、正解データ）については、音楽大学の大学院生に作成していただいたものを使用した。ある主旋律中の音に対して生成された副旋律データが、正解データの音と完全一致した場合、正しい副旋律が生成されたとする。ただし実際には正解となる副旋律は複数存在する。そのため、作成していただいた正解データは複数ある正解副旋律のうちの1つであり、主旋律と副旋律が完全一致しなかった場合でも、生成した副旋律が音楽的な不正解であるとは限らない。

4.2 評価結果

HMM とルールベースの2つの手法でそれぞれ上でハモる場合と下でハモる場合で副旋律を生成し、それぞれ評価を行った。副旋律と正解データの一致率（以下精度とする）を表4.1、各手法で生成した副旋律の85曲全体の精度を図4.1、図4.2、図4.3、図4.4に示す。

表 4.1: 副旋律生成結果の精度

	HMM			ルールベース		
	平均	最大	最小	平均	最大	最小
上	52.3%	64.5%	24.3%	55.5%	75.0%	33.6%
下	53.6%	71.0%	30.5%	64.2%	87.9%	37.6%

HMM(上)

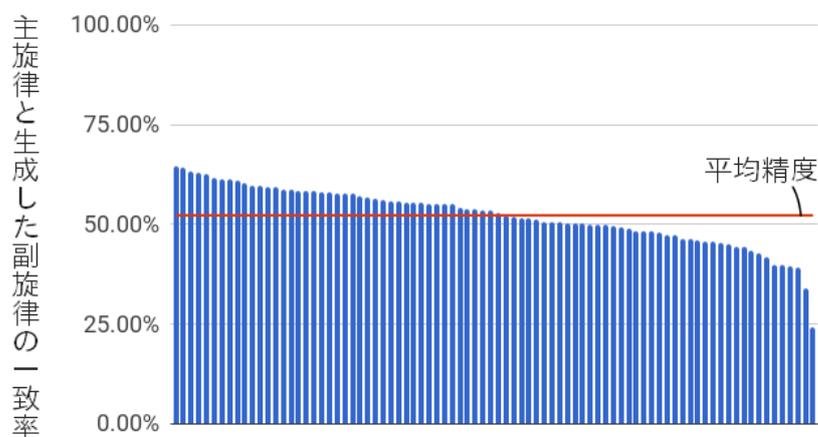


図 4.1: HMM で生成した上ハモリの 85 曲全体の精度

ルールベース(上)

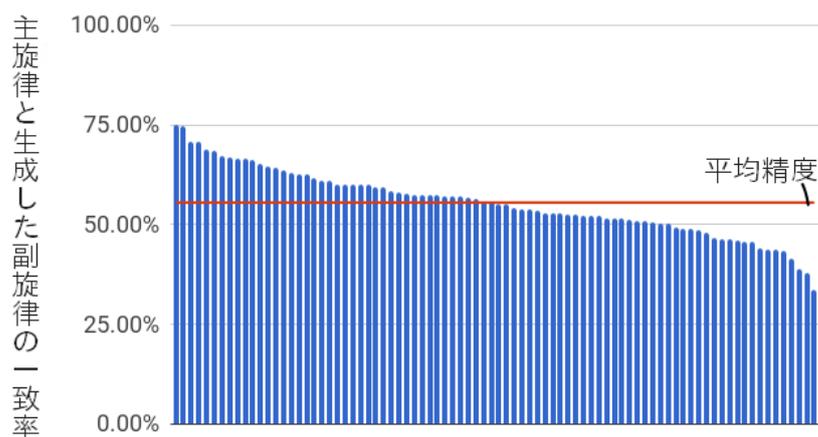


図 4.2: ルールベースで生成した上ハモリの 85 曲全体の精度

HMM(下)

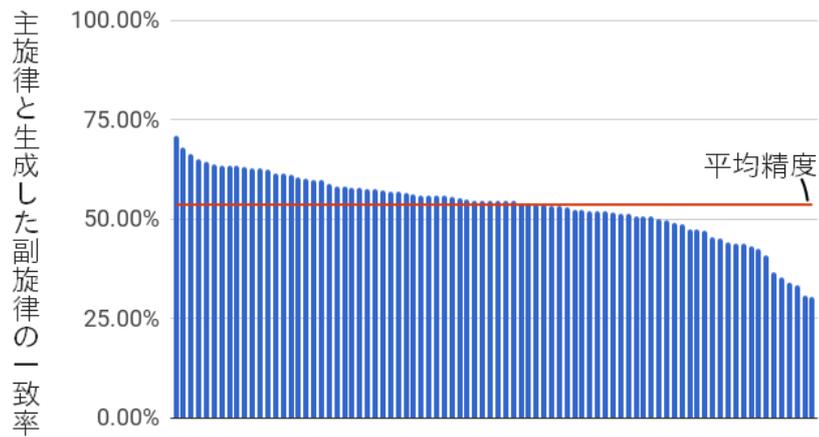


図 4.3: HMM で生成した下ハモリの 85 曲全体の精度

ルールベース(下)

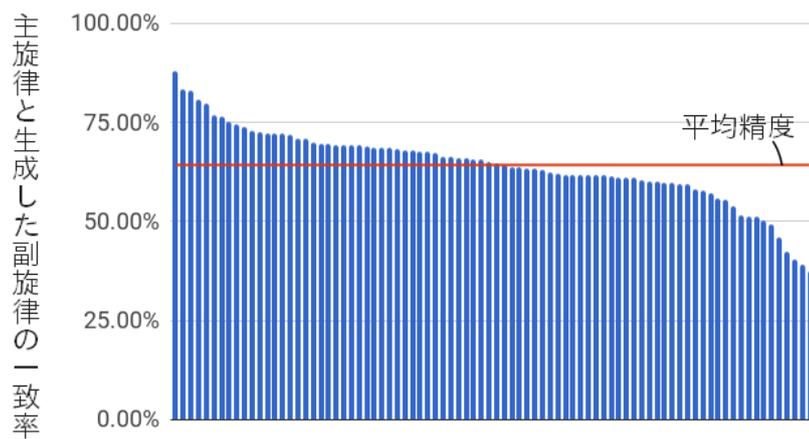


図 4.4: ルールベースで生成した下ハモリの 85 曲全体の精度

副旋律を主旋律よりも高くする（上でハモる）場合と副旋律を主旋律よりも低くする（下でハモる）場合ともに、HMM で生成した副旋律よりもルールベースで生成した副旋律の方が平均精度が高い結果となった。

4.3 考察

4.3.1 副旋律評価

HMM で生成した副旋律よりルールベースで生成した副旋律の方が平均精度が高い結果となった理由について考察する。まず第1に、正解データに含まれているユニゾンが少ないにも関わらず、HMM で生成した副旋律には、ユニゾンが多く含まれていたことが挙げられる。正解データ中に含まれるユニゾンの割合は、全曲を平均して0.20%であった。一方、HMM で生成した副旋律では4.78%であった。そのうち、正解データと一致するユニゾンを生成できた割合は1.22%であった。このことから、HMM で生成した副旋律がユニゾンだった部分で大きく精度が下がってしまったことが分かる。また、正解データ中に含まれるユニゾンの特徴として、これらは休符前後に多く見られる傾向にあったが、HMM で生成した副旋律は休符の有無に関わらずユニゾンを生成していたため、精度が下がってしまったと考えられる。ルールベースで生成する副旋律では、ユニゾンが生成されないが、正解データ中のユニゾンの割合が少ないため、精度への影響が少なかったと考えられる。

第2に、HMM で生成した副旋律は、主旋律の音名の前後関係による影響が大きかったことがあげられる。図4.5, 図4.6 はそれぞれ、「上でハモる」場合、「下でハモる」場合で正解データ、ルールベース、HMM で生成した副旋律のうち、連続で同じ音に遷移した回数を表す。ここから、HMM で生成した副旋律は、特に連続で同じ音に遷移する回数が多く、本来副旋律中に高低差をつけなければいけない部分でも、同じ音が連続で続くことになってしまったため、精度が下がったと考えられる。

HMM で生成した副旋律が、同じ音に遷移する回数が多くなってしまった

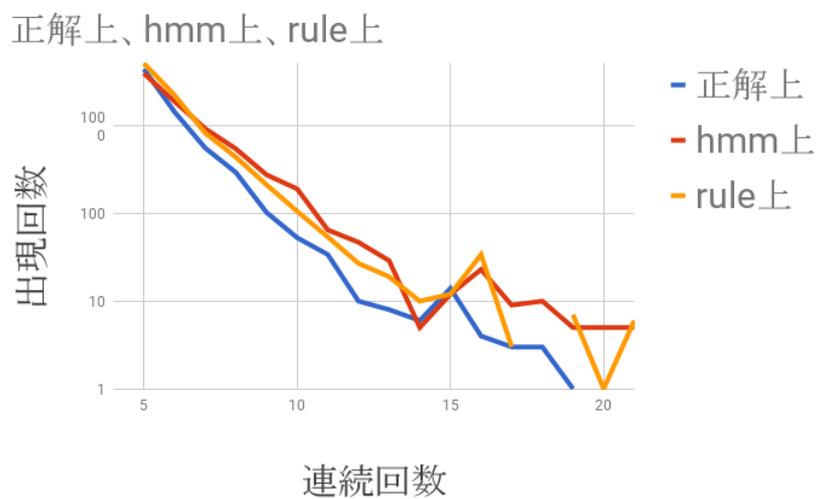


図 4.5: 85 曲中で同じ音名が連続する箇所の出現回数 (「上で八モる」場合)

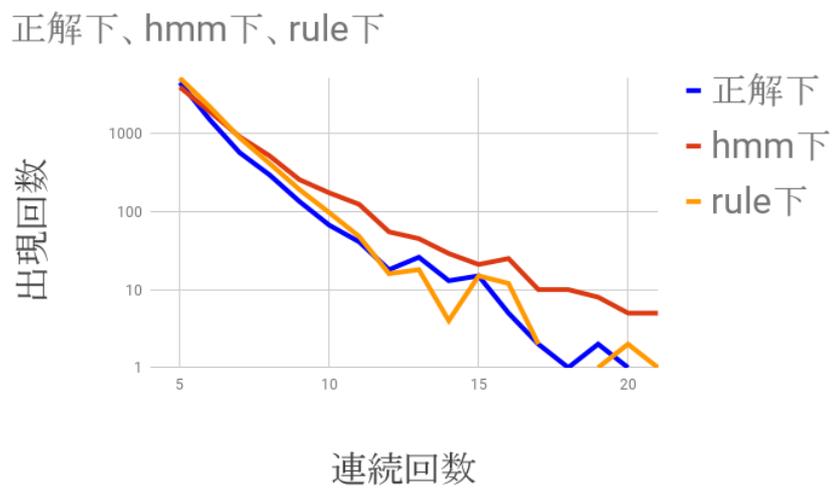


図 4.6: 85 曲中で同じ音名が連続する箇所の出現回数 (「下で八モる」場合)

理由としては、利用する遷移確率の影響があると考えられる。同じ音が10音以上連続で続いた30曲の遷移確率に着目すると、「下で八モる」場合では、 $A\sharp$ を除いた11種類の音名において、「上で八モる」場合では、 $C\sharp, F$ を除いた10種類の音名において、同じ音名に遷移する確率が他の音名に遷移する確率より高かった。一方、それ以外の55曲の遷移確率に着目すると、「下で八モる」場合では、 $D\sharp, E, A, A\sharp, B$ の5種類の音名において、「上で八モる」場合では、 $C\sharp, D, D\sharp, E, G, G\sharp$ の6種類の音名において、同じ音名に遷移する確率が他の音名に遷移する確率より高かった。これらのことから、HMMで副旋律を生成するにあたって、主旋律の音名の前後関係から算出した遷移確率の影響が大きかったためだと考えられる。

4.3.2 生成された副旋律の主観評価

今回用意した正解データはある主旋律に対する理想的な副旋律の1例であり、正解データと一致しなかったからといって、音楽的に不適切であるとは限らない。そのため、4.2節で行った副旋律評価でHMMで生成した副旋律の精度が35%以下となった6曲(表4.2)、ルールベースで生成した副旋律の精度が40%以下となった5曲(表4.3)を別の音楽家に評価してもらった。評価方法は、副旋律の歌いやすさは考慮せず、生成した副旋律が音楽的に妥当であるかを評価することとする。評価結果を表4.4、表4.5に示す。HMM、ルールベースで生成した副旋律の評価は、4.2節にて行った実験の精度を上回りいずれも67%を超える評価となった。このことから、HMM、ルールベースで生成した副旋律は、一定の音楽的妥当性があることが分かった。

表 4.2: 再評価をする HMM で生成した副旋律の精度

曲名 (上下ハモリ)	正解データとの完全一致率
あったかいんだからぁ (上)	34.14%
あったかいんだからぁ (下)	30.51%
花 (下)	33.45%
夏祭り (上)	24.36%
366 日 (下)	34.09%
大阪 LOVER (下)	30.98%

表 4.3: 再評価するルールベースで生成した副旋律の精度

曲名 (上下ハモリ)	正解データとの完全一致率
あったかいんだからぁ (上)	38.98%
あったかいんだからぁ (下)	39.23%
花 (下)	37.56%
青いベンチ (上)	37.77%
夏祭り (下)	33.64%

表 4.4: HMM で生成した副旋律の再評価結果

曲名 (上下八モリ)	専門家による評価
あったかいんだからぁ (上)	73.51%
あったかいんだからぁ (下)	83.24%
花 (下)	97.46%
夏祭り (上)	81.58%
366 日 (下)	79.81%
大阪 LOVER (下)	73.80%

表 4.5: ルールベースで生成した副旋律の再評価結果

曲名 (上下八モリ)	専門家による評価
あったかいんだからぁ (上)	67.03%
あったかいんだからぁ (下)	78.92%
花 (下)	90.30%
青いベンチ (上)	85.33%
夏祭り (下)	85.53%

第5章 副旋律歌唱練習システムへの応用

主旋律を歌唱する相手の特性にあわせて副旋律を練習を行うためには、副旋律を練習しているユーザがシステム上で、直接主旋律歌唱者の特徴を再現できることが必要である。本研究では、歌唱相手の再現すべき特徴の内の1つであるキーに着目し、システム上で調整を行えるようにする。

主旋律につられない練習を行えるようにするために、ユーザが自身の歌声をよく聴くこと（聴覚的アプローチ）[10]、ユーザが自身の音高と副旋律が一致しているかを確認すること（視覚的アプローチ）が必要である。これらのことから、聴覚的アプローチではユーザが自らで発した歌声を自身で聞こえるようにする。視覚的アプローチでは、ユーザが自らで発した歌声の音高を視覚的に表示し、理解できるようにする。以上2点のアプローチを満たす環境で練習を行うことで、主旋律につられない練習ができると考えられる。

図 5.1, 図 5.2 が本研究で実装したシステムである。

図 5.1 は楽曲の選択画面で、図 5.2 は練習画面の様子である。

5.1 システムの流れ

1. 練習を行いたい楽曲を選択する。
2. 練習を行う楽曲を再生する際の様々な設定を行う。
 - 主旋律, 伴奏の音量を選択する。
 - 主旋律のキーを選択する。
 - 上八モリ, 下八モリの選択をする。



図 5.1: 練習曲を選択し，練習モード，主旋律のキー，伴奏や主旋律の音量，ハモリの上下を設定を行う．



図 5.2: 画面上部で主旋律，副旋律，ユーザの音高を表示をしている．ユーザの歌唱した音高をオクターブ上下して表示することもできる．画面中央で歌詞を表示している．画面下部では，はじめから，早送り，巻き戻しをして再生を行えるボタンがある．

3. 1 で選択した楽曲の MIDI データから楽曲の主旋律と伴奏の情報を取得し、それに基づいて副旋律を生成する。
4. 2 で選択した設定を 3 で生成した副旋律、主旋律、伴奏にそれぞれ適用し、再生を行う。
5. 歌唱練習中のシステムの動きとユーザの操作については以下の通りである。
 - ユーザの歌唱データを 5ms 毎に保存し、基本周波数推定法 DIO[11] によって F0 推定を行う。
 - 画面上部に主旋律、副旋律、ユーザ自身の音高 (F0) をリアルタイムにそれぞれ表示する。
 - 上下の方向キーを押すことで、ユーザ自身の音高表示を行う際、オクターブ上げた (下げた) 状態で表示を行うことができる。
 - はじめからボタンを押すと曲がはじめから再生される。巻き戻しボタンを押すと 4 小節前が再生される。早送りボタンを押すと 4 小節先が再生される。

5.2 相手の特性 (キー) に合わせた練習

相手の特性 (キー) に合わせた練習では、自動生成した副旋律を、半音上げ下げする事で楽曲のキーの変更を行うことを可能にした。

5.3 主旋律につられない練習

副旋律練習者には、自身の音高を理解できる環境で練習を行う。理解の方法として視覚と聴覚による 2 つのフィードバックを行う。

視覚的なフィードバックでは、ユーザが歌唱した音高が図 5.3 のように表示され、ユーザは目標である副旋律の音高を目指し練習を行う。

聴覚的なフィードバックでは、ユーザが歌唱した声をそのままヘッドフォンに出力することで行う。自分の歌声がヘッドフォンに返ってくる機能（以降ループバック機能とする）によってユーザ各人が自分の返ってくる歌声の音量を調節することで、実際のカラオケの現場でもつられないような練習をすることを可能にする。具体的には、練習前半ではヘッドフォンから聞こえる自分の歌声が大きい状況で練習してもらい、自分の歌声をよく理解できる状況で練習を行ってもらう。練習後半では、徐々にヘッドフォンから聞こえる自分の歌声を小さくして練習してもらい、実際のカラオケのように自分の歌声が理解しづらい状況で練習を行ってもらう。この手順で練習を行う事で、実際のカラオケでもつられないようになるための練習を行う。

5.4 その他の機能

上記の実装の他に実装した機能を 4 つ紹介する。

練習楽曲の主旋律を理解していない状況では、副旋律との関係（どのくらい音が離れているか）を理解することが難しい。そのため図 5.1 のように主旋律のみを練習できる機能を実装した。

伴奏が大きく、主旋律や副旋律の音が聞き取りにくくなることを防ぐために、図 5.1 のように伴奏の音量を調節できる機能を実装した。同様に、主旋律の音量が大きく、副旋律が聞き取りにくくなることを防ぐために、図 5.1 のように主旋律の音量調節ができる機能を実装した。さらに、伴奏や主旋律の音量調節を行えるようにしたことで、まずは主旋律、伴奏がほとんど聞こえない状態から練習を開始し、段階的にそれぞれの音量を上げ、実際に本番で歌唱するシーンに徐々に近づけながら練習することが可能になる。

ユーザの音域によっては原曲の高さに合わせて歌うことができないことがある。そのため図 5.2 のように、ユーザの発声した歌声の音高表示を視覚的に行う際、実際の音よりもオクターブ上げた（もしくは下げた）状態で表示できるようにする機能も追加した。これによって、ユーザは原曲に対してオクターブ上げた（もしくは下げた）状態で歌唱を行っても、原曲のキーを歌う

感覚で練習できているかを確認できる。繰り返し練習がしやすいように、図 5.2 のように、はじめから始める機能，早送り・巻き戻し機能を実装した。



図 5.3: 画面上部にユーザの歌唱した音高が表示されている。ユーザは赤で表示されている目標の副旋律の音高に近づけるように練習を行う。

第6章 結 論

本稿は、ユーザが副旋律を練習する際に生じる、“副旋律を担当するユーザが副旋律を入手することが困難”、“主旋律歌唱者の特性(キー)を考慮できない”、“主旋律につられてしまう”の3つの問題の解消を狙ったものである。副旋律を担当するユーザが副旋律を入手することが困難であるという問題については副旋律の自動生成をHMM, ルールベースの2手法で実現し解決した。主旋律歌唱者の特性(キー)を考慮できない問題、主旋律につられてしまう問題については自動生成した副旋律を利用した副旋律練習システムを実装し、システムの機能として練習楽曲の主旋律、副旋律、伴奏の音高を変更できる機能と自身の音高の視覚的フィードバックができる機能を実装することによってそれぞれ解決した。

また、自動生成した副旋律が、どの程度適切であったかをHMMで生成した場合とルールベースで生成した場合で検証した。使用した正解データは音楽大学の大学院生が作成した副旋律を使用した。上で八モる場合、下で八モる場合で練習を行ったところ、上で八モる場合の平均精度はHMMで生成した副旋律は52.3%、ルールベースで生成した副旋律は55.5%となった。下で八モる場合の平均精度はHMMで生成した副旋律は53.6%、ルールベースで生成した副旋律は64.2%となった。上で八モる場合、下で八モる場合の両方でHMMで生成した副旋律よりもルールベースで生成した副旋律の方が平均精度が高い結果となった。その原因は2つあり、正解データに含まれているユニゾンが少ないにも関わらず、HMMで生成した副旋律には、ユニゾンが多く含まれていたことと、HMMで生成した副旋律は、主旋律の音名の前後関係による影響が大きかったことが挙げられる。

前述の副旋律の検証における各手法で生成した副旋律の一致率は正解デー

タとの比較であるので、一致しなかった場合でも音楽的に不適切であるとは限らない。このことから、HMM で生成した副旋律の精度が 35%以下となった 6 曲，ルールベースで生成した副旋律の精度が 40%以下となった 5 曲の計 11 曲に対して音楽的に適切であるかの評価を別の音楽の専門家に依頼した。その結果，11 曲全曲において 67%を超える評価となったため，HMM，ルールベースで生成した副旋律は一定の音楽的妥当性があることが分かった。

今後は，HMM での生成した副旋律の精度向上や実装した副旋律練習システムを利用して，ユーザが副旋律を正しく練習することができるか検証していく予定である。また，実装した副旋律練習システムを利用して，ユーザが副旋律を正しく練習することができるか検証していく予定である。

参考文献

- [1] 栗原拓也, 横溝有希子, 竹腰美夏, 馬場哲晃, 北原鉄朗, スマートタンバリン: 音と光で場を盛り上げるカラオケ支援システム, 研究報告音楽情報科学 (MUS), Vol. 2017, No. 3, pp. 1-5, 2017.
- [2] 浦川 雄一, 勝頼郁代, 川島達也, はずれた音程を修正するカラオケシステムー合成した音声の歌声らしさの向上ー, 平成 21 年度電気関係学会九州支部連合大会予稿集, 03-2P-02, 2008.
- [3] 池内友次郎, 長谷川良夫, 石桁真礼生, 松本民之助, 島岡譲, 柏木俊夫, 丸田昭三, 小林秀雄, 三善晃, 未吉保雄, 佐藤真, 和声理論と実習, 音楽之友社, 1964.
- [4] エバンズベンジャミンルカ, 棟方渚, 小野哲雄, 和声学に基づく合唱譜自動作成システム, 和声学に基づく合唱譜自動作成システム "CMY " 転回系の実装と評価, 情報処理学会研究報告, Vol. 2013, No. 37, pp. 1-4, 2013.
- [5] 三浦雅展, 青山容子, 谷口光, 青井照博, 尾花充, 柳田益造, ポップス系の旋律に対する和声付与システム: AMOR, 情報処理学会論文誌, Vol. 46, No. 5, pp. 1176-1187, 2005.
- [6] Syunpei Suzuki, Tetsuro Kitahara, Four-part Harmonization Using Bayesian Networks: Pros and Cons of Introducing Chord Nodes, *Journal of New Music Research*, Vol. 43, No. 3, pp. 331-353, 2014.

- [7] 菅原 啓太, 西本卓也, 嵯峨山茂樹, HMM と音符連鎖確率を用いた旋律への自動和声付け, 情報処理学会研究報告音楽情報科学 (MUS), Vol.2003, No. 127, pp. 49–54, 2003.
- [8] 福本愛由星, 橋田光代, 片寄晴弘, RelaPitch: 歌って鍛える相対音感習得支援システム, エンタテインメントコンピューティングシンポジウム 2016 論文集, Vol. 2016, pp. 332–333, 2016.
- [9] Tomoyasu Nakano, Masataka Goto, Yuzuru Hiraga, MiruSinger: A singing skill visualization interface using real-time feedback and music CD recordings as referential data, *Ninth IEEE International Symposium on Multimedia 2007 - Workshops*, pp. 75–76, 2007.
- [10] 小原孝, うまく歌える!演奏がうまくなる!楽譜が読める!音感スーパーレッスン, ナツメ社, 2012.
- [11] 森勢将雅, 河原英紀, 西浦敬信, 基本波検出に基づく高 SNR の音声を対象とした高速な F0 推定法, 電子情報通信学会 論文誌 D, Vol. J93-D, No. 2, pp. 109–117, 2010.

謝 辞

本論文を作成するにあたり、指導教員の北原鉄朗准教授から、丁寧かつ熱心なご指導を賜りました。的確な指導や多くの議論によって、自身の研究がより良いものとなりました。ここに感謝の意を表します。

日常の議論を通じて気づかなかった部分や改善点、客観的視点からの評価等を頂いた北原研究室の先輩である大野先輩、棚橋先輩に深く感謝致します。共同研究者の小笠原梢氏には、自身の作成したシステム実験について担当して頂きました。難しい部分であったと感じますが、丁寧に結果、考察等まとめて頂きました。本当にありがとうございます。

また、副旋律の正解データを85曲分作成して頂いた宮下和也氏、生成した副旋律の再評価を快く受けて下さった橋田光代先生、被験者実験にご協力頂いた方々、切磋琢磨し研究に励んだ北原研究室の同期、後輩の皆様には感謝致します。