

## ハミングによる単旋律の自動採譜

清水純†, 丸山剛志†, 三浦雅展‡, 柳田益造†

†同志社大学 工学部,

‡龍谷大学 理工学部

本研究では、単旋律のハミングによる歌唱の自動採譜を目標としている。音高については、その絶対値は正しくなくても、相対的に正確に歌えていれば各階名の周波数比は保たれることを利用し、平均律音階の周波数比と等しい間隔で並んだくし型のテンプレート(鍵盤テンプレート)を用いたマッチングにより、利用者にとっての音階位置を認識し、階名を推定する。また、音価についても、音符の音価が基本的には2のべき乗の体系をしていることを利用し、2のべき乗間隔で並んだくし型のテンプレート(音価テンプレート)を用いて、オンセット間隔の局所的頻度分布とのマッチングをとることにより、テンポが揺らいでも音価を推定できる方法を提案する。

## Automatic Scoring of Melodies Sung by Humming

Jun SHIMIZU†, Takeshi MARUYAMA†, Masanobu MIURA‡, Masuzo YANAGIDA†

†Faculty of Engineering, Doshisha Univ., Kyoto

‡Faculty of Science and Technology, Ryukoku Univ., Shiga

The objective of this research is automatic scoring of melodies sung by humming. The absolute frequency position of a musical scale for a singer is estimated by matching a “keyboard template” having line spectra at frequency points corresponding to the well-tempered scale on a key of variable frequency with occurrence distribution of fundamental frequencies. The syllable names can be estimated based on the scale determined by the method described above. A “note-value template” having line components at 2's power interval is used to detect the unit of note value by finding accordance between the note-value template of a variable time unit with occurrence distribution of IOI(Inter-Onset Interval). The method is valid even if the tempo fluctuates by taking the range of evaluation to be localized.

### 1. まえがき

本研究の目標は、ハミングによる歌唱の自動採譜である。自動採譜は、音になった音楽の記譜、民族音楽に関するデータベースのコンテンツ作成等の効率化に有用な技術であり、また、音(旋律の歌唱や演奏、録音された音楽の再生音)からの楽曲検索、素人を対象とした作曲・編曲補助などにも用いることができる。その他にも、携帯電話端末からメロディーを音符で入力すると、それに和音を付与するシステム[1]があるが、携帯電話端末からの音符入力インターフェースの向上にも役立つと考えられる。

これまでに、自動採譜に関する研究としては、知識処理的アプローチによるピアノ・ギター・三味線の自動採譜[2]、クラスタリング手法を用いたバイオリン二重奏の採譜[3]、弦楽四重奏の採譜[4]などがあ

るが、楽器による単旋律を対象とした場合はほぼ実用レベルに達しており、また、単一楽器によるポリフォニーに対しては、楽器種を限定し、同時発音数などに関して強い制約を設けても構わない場合は、人間による採譜の労力を少しは軽減できるレベルにあると言える。

主に楽器音を対象とした自動採譜システムに歌唱を分析させると、それなりの結果は得られるが、歌唱は楽器での演奏とは違い、基準音高を与えられるか、歌い手が絶対音感をもっていなければ、音高の絶対位置が正規の高さからずれてしまう可能性があることや、各音の立上り、立下りが不明瞭になり、定常部以外の部分音高が不安定であることや、テンポが揺れることなど、自動採譜を困難にする要因がいくつも存在する。

音高の絶対位置がずれてしまうことが原因で、正規の音高基準で音高判定を行ったのでは、正確な階名の推定が行えない場合がある。また、歌唱を対象とした自動採譜を行う市販ソフト[5]で、音価の抽出を正確にするために、入力時に指定のテンポで歌う必要があるものがあるが、その条件の下でも、音価推定の動作性能はあまりよくない。

本研究では、音階の絶対位置のずれを許容し、かつ、歌唱のテンポが変化しても、安定に自動採譜を行うことができる階名・音価の推定法を提案する。階名・音価ともに調性音楽の構造に基づいたテンプレートとの整合性を見ることによって推定する。

階名の推定方法としては、音階の絶対位置は標準の値でなくても相対的に正確に歌っていれば、各階名間の周波数比はその音程によって一定であることを利用する。その周波数比に等しい間隔で並んだくし型の周波数テンプレートを用意し、基本周波数の頻度分布との最も整合の取れる位置へ音階をもってくことにより、入力した人にとっての音階位置と主音の周波数を推定し、それをもとに各歌唱音の階名の推定を行う。

調性音楽においては、音価は2のべき乗の体系をしていることを利用し、単位時間を可変として2のべき乗間隔で並んだくし型の時間テンプレートを用意し、歌唱に含まれる隣接2音のオンセット間隔(IOI, Inter-Onset Interval)の頻度分布との最大整合点を見つけることによって、歌唱のテンポ(単位音価の絶対時間長)を推定する。頻度分布をとる範囲をどのようにするかで、局所的テンポも長い区間の平均テンポも自由に推定できる。この考え方によって、歌唱のテンポによらず、音価の推定を適応的に行うことができる。

2章では、上述の階名・音価の推定に基づいた採譜の方法を詳しく述べ、3章では、その方法を具体的にどのように実現しているかを述べる。

## 2. 採譜の方法

### 2.1 各音の時間区間の検出

単旋律の歌唱では、ある時刻においては、1つの音しか発せられていない(壁などからの反射はここでは考えない)。従って、単旋律の歌唱に対して採譜を行うには、まず、ほぼ一定の音高で発声されている時間区間に区切る必要がある。言い換えると、譜面上での各音符に対応する時間区間を調べるのである。そして、その時間区間に対応する音価を推定し、さらには、その時間区間の音高情報から階名を推定すればよい。

そこで、まず、1つの音に対する時間区間の検出

方法を説明する。歌唱では、ある音と音の間に一般には(スラーやタイがかかっていない場合)無音の時間区間が生じる。これは duty ratio(音価中で実際に音が発されている時間の割合、ピアノではダンパーがはたらくまでの、つまり、キーを離すまでの時間の割合)が1以下であることによるもので、無音区間はその前の音がもつ音価に含まれるものとみなす。従って、1つの音に対する時間区間を検出するには、それぞれの音の立上り時刻を検出し、次の音の立上り時刻までの時間(オンセット間隔:IOI)を求め、このIOIを1つの音に対する時間区間と考えればよい。ここでは、休符は認識せず、休符の音価はその前の音の音価に含まれるものとみなす。これは、歌唱での休符はその開始時点の推定がほとんど不可能であるためである。

音の立上り時刻を検出する方法としては、大きく2つの方法が考えられる。1つは、波形の振幅に対して、閾値処理を行うものである。もう1つは、入力信号の音高情報を時間方向に並べ、音高が大きく変化するところを音の立上りとみなす方法である。歌唱は楽器音とは違い、その振幅が不安定で、極端に大きな声で歌う部分や、逆に小さな声で歌う部分が存在することがある。従って、振幅に基準を置いた方法では、正確な検出が困難であろうと考え、ここでは、音高情報を用いる方法を採用する。

### 2.2 音価の推定方法

本システムでは、利用者に対してある一定のテンポを指定して歌わせるのではなく、利用者が自由なテンポで歌うことを許容する。

歌唱のテンポは、指示がなければ、かなりの自由度がある。従って、歌唱中の各音の音価を推定するためには、その歌唱における、四分音符、八分音符などの各音符の音価に対応するIOIを調べる必要がある。これが分かれば、各音に対するIOIが、どの音価に最も近いかを調べることにより、各音の音価を推定することができる。調性音楽では、音価は基本的には2のべき乗の体系をしているので、歌唱の中で最も短い音符の音価に対応するIOIを調べることができれば、その他の音符の音価に対応するIOIも分かったことになる。本稿では、歌唱の中で最も短い音符の音価に対応するIOIを「基準IOI」と呼び、 $\tau$ で表すこととする。他の音価は、ほぼこれの2のべき乗の長さをとることになる。

基準IOIを調べるために、歌唱中のIOIの頻度分布をとる。すると、複数のピークが現れることになり、それらのピークの中で、2のべき乗系列でのIOIが最も短いもののIOIが基準IOIであると考えらるこ

とができる．ここで，音価が 2 のべき乗体系をなすことを考慮して，図 1 のような形の関数  $T_i(t, \tau)$  をテンプレートとして用意する．このテンプレートを「音価テンプレート」と呼ぶこととする．

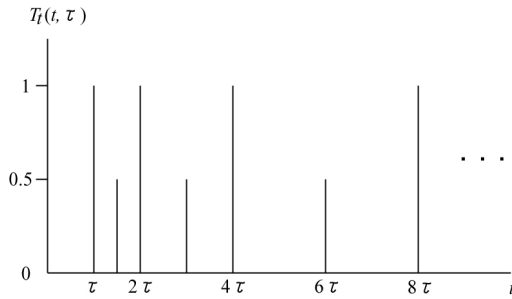


図 1 関数  $T_i(t, \tau)$

先にも述べたが，八分音符，四分音符などの基本の音価は 2 のべき乗の体系をしているので， $t$  が  $\tau$  の 2 のべき乗倍となるとき，関数値 1 を与えている．また，それらの音符の 1.5 倍の長さをもつ，付点の付いた音符も存在する．この付点の付いた音符を，基本の音符であると誤認識してしまうことを防ぐために， $t$  が  $\tau$  の 2 のべき乗の 1.5 倍となるとき，基本系よりも重みを抑えた 0.5 という関数値を与えている．

関数  $T_i(t, \tau)$  は 2 つの時間  $t$  と  $\tau$  を変数としてもつ関数で， $\tau$  は基準 IOI を表す．また， $T_i(t, \tau)$  は  $\tau$  によって時間方向に伸縮する関数である． $\tau$  を変化させながら，IOI の頻度分布  $F_i(t)$  と  $T_i(t, \tau)$  の  $t$  軸上での整合度を調べ，整合度が最大となるとき  $\tau$  を，その歌唱における基準 IOI であると推定することができる．

このようにして求めた基準 IOI を用いて，各音の IOI が基準 IOI の約何倍であるか(基本音符なら 2 のべき乗倍，付点音符なら基本音価の 1.5 倍)を調べることにより，各音の音価を推定することができる．例えば，ある歌唱の中で最も短い音符は十六分音符であり，ある音に対する IOI が基準 IOI の約 4 倍であったとすると，その音は四分音符の音価を持つと推定することができる．なお，歌唱のテンポは揺れるので，緩急の大きい歌唱の基準 IOI の推定は，局所的に行うことが望ましい．IOI の頻度分布を局所的にとった場合には，サンプルが少なくなるが，そのような場合でも，音価テンプレートはうまく機能してくれることが期待できる．

### 2.3 階名の推定方法

1 つの音と判定される時間区間内での，基本周波数の平均値を求め，それがどの階名の基準周波数に最も近いかを調べることにより，階名の推定を行う．従来の音高推定システムでは，標準位置での音階の

基準周波数(A4=440 or 442Hz)を用いることが多いが，それでは任意の基準周波数で歌うことを許した場合に対してうまく推定を行えない．

絶対音感を持たない人は，歌い始めの高さを与えられないと，譜面上の調の正確な高さで歌うことが困難であるが，相対的に正確に歌えていれば，絶対的な音の高さは正規の高さでなくても，正確に歌えていると考えるべきであり，特に鼻歌のように気安く歌う場合に，絶対的な音の高さを利用者に要求することは不可能であるので，その問題はシステム側で対処すべきである．

仮に，絶対位置での音階の基準周波数を用いて，階名の推定を行ったとする．そして，利用者は八長調のドの音を発したつもりが，C と C# のほぼ中央の周波数の音を発してしまったとする，一般的に，いわゆる音痴の人でない限り，歌唱では音程関係はほぼ規定通りに保たれるので，その利用者にとっての D の音は，D と D# のほぼ中央の周波数，E の音は E と F のほぼ中央の周波数，といったように，全ての音が絶対位置での音階の基準周波数から一定の比でずれてしまう．すなわち，全ての音に対して，一定の比による補正を行わなければ，正確な階名の推定が困難となってしまう．

このような状況に陥らないため，本研究では，音階の基準周波数を絶対位置に決めてしまうのではなく，入力音声の基準周波数の頻度分布をもとに，利用者にとっての音階の基準周波数を求め，それをもとに階名を推定する方法をとる．

まず，入力音声全体における，基本周波数の頻度分布をとると，全音，または，半音間隔でピークが現れると考えられる．そこで，図 2 のような形の関数  $T_f(f, f_c)$  をテンプレートとして用意する．このテンプレートを「鍵盤テンプレート」と呼ぶこととする．

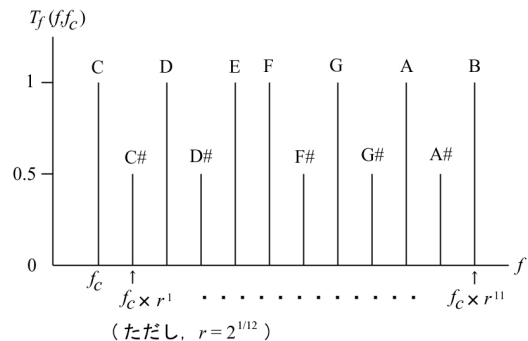


図 2 関数  $T_f(f, f_c)$

関数  $T_f(f, f_c)$  は，周波数  $f$  と C の基準周波数に対応する可変基準周波数  $f_c$  を変数としてもつ関数である．図 2 の横軸は対数軸で， $T_f(f, f_c)$  は  $f_c$  の値によ

って  $f$  軸方向に平行移動するような形の関数であり、ピアノの白鍵に対応する相対音高で関数値 1、黒鍵に対応する相対音高で関数値 0.5、その他で 0 となる、くし型の関数である。

$f_c$  の値を変化させながら、基本周波数の頻度分布との整合度を調べ、整合度が最大となるときの  $f_c$  を、その利用者にとっての(移動ドによる)ドの基準周波数と推定することができる。

また、ここでは各階名間の周波数比は平均律に保たれると仮定し、ド#, レ, レ#, ミ...と、全ての階名の基準周波数を求めることができる。

このようにして求めた、利用者にとっての音階の基準周波数  $f_c$  をもとに、各音の周波数平均がどの階名の基準周波数に最も近いかを調べることにより、各音の階名を推定する。

この方法は、白鍵をベースとした、ヒューリスティックな規則をもとに関数  $T_f(f, f_c)$  を定義しているため、八長調以外の調の曲を歌った場合、移動ドとしての階名を推定することになる。もし、絶対的な音の高さが必要であれば、出力の際に元の高さに近い所の規定値に設定し直せばよいと考えているため、本システムでは、とりあえずは、移動ドとしての階名をそのまま出力することとする。

### 3. 処理の手順

#### 3.1 歌唱の入力

マイクrohonを用い、ハミングによる歌声を入力する。収録は、環境雑音が混入しないように、接話マイクによるか、あるいは防音室で行うことが望ましい。

#### 3.2 各音に対する IOI の検出

入力音声に対して、ケプストラムを用いて、基本周波数を抽出する。入力音声のサンプリング周波数は 22.05KHz で、分析フレームを 256 サンプル点ずつずらしながら各分析フレームにおける基本周波数を抽出する。あるフレーム  $t$  における基本周波数の値をセント値で表現したものを  $F0(t)$  とおく。変数  $t$  は分析フレーム番号である  $F0(t-1)$  と  $F0(t)$  との差が、50 セント(半音の半分)以上、つまり、式(1)を満たすときの  $t$  を基本周波数の立上り時刻(実際はフレーム位置)とする。

$$|F0(t) - F0(t-1)| \geq 50 \quad (1)$$

この判定基準で、ある音の立上り時刻から次の立上り時刻までの時間(IOI)を求める。判定基準からわかるように、ここではタイがかかった同音は連続した音符とみなし、トリルの各音は個別の音とみなす

ことになる。

### 3.3 音価の推定

前節で求めた IOI の頻度分布  $F_i(t)$  をとる。

関数  $T_i(t, \tau)$  を以下の式により定義する。

$$T_i(t, \tau) = \begin{cases} 1 & \text{for } t = 2^n \tau (n=0,1,2 \dots) \\ 0.5 & \text{for } t = 1.5 * 2^n \tau (n=0,1,2 \dots) \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

関数  $T_i(t, \tau)$  と、関数  $F_i(t)$  との時間軸上の整合度  $C_i(\tau)$  を以下の式により求め、 $C_i(\tau)$  の最大値を与える  $\tau = \tau_0$  を基準 IOI とする。 $\tau$  の変化幅はサンプリング周期とする。

$$C_i(\tau_0) = \max C_i(\tau) \quad (3)$$

$$\text{ただし, } C_i(\tau) = \int T_i(t, \tau) \cdot F_i(t) dt \quad (4)$$

各音に対応する IOI が  $\tau_0$  の約何倍かを調べることにより、音価の推定を行う。ある音に対する IOI を  $T$  とすると、

$$A(m_0) = \min A(m) \quad (5)$$

$$(m = 1, 1.5, 2, 3, 4, 6, 8, 12, 16, \dots)$$

$$\text{ただし, } A(m) = \left| \frac{T}{\tau_0} - m \right| \quad (6)$$

$$\begin{cases} m_0 = 1: \text{八分音符の音価} \\ m_0 = 1.5: \text{付点八分音符} \\ m_0 = 2: \text{四分音符の音価} \\ m_0 = 3: \text{付点四分音符の音価} \\ m_0 = 4: \text{二分音符の音価} \\ m_0 = 6: \text{付点二部音符} \\ \vdots \end{cases} \quad (7)$$

となる。

式(5)を満たす  $m_0$  を見つけ、式(7)に従って音価を推定する。 $m_0=5$  や、 $m_0=7$  といった音価は音楽的には存在しないが、本システムでは、MIDI ファイルに出力する際、基準の音価(= $\tau_0$ )の  $m_0$  倍の音価としてそのまま出力する。例えば、付点四分音符の後に四分休符がある場合、ここでは休符は認識しないので、 $m_0=5$  という結果となる。あるいは、音価の認識が正しくできていないときに、このような音楽的には存在しない音価が出力される。

現状では、対象を童謡にしているの、最も短い音符を八分音符と設定してあるが、16分音符、あるいは32分音符に設定することも可能である。

### 3.4 階名の推定

入力音声全体における，基本周波数の頻度分布  $F_f(f)$  をとる．

関数  $T_f(f, f_c)$  を以下の式により定義する．

$$T_f(f, f_c) = \begin{cases} 1 & \text{for } f = f_c \times r^k (k = 0, 2, 4, 5, 7, 9, 11) \\ 0.5 & \text{for } f = f_c \times r^k (k = 1, 3, 6, 8, 10) \\ 0 & \text{otherwise} \end{cases} \quad (8)$$

ただし， $r = 2^{\frac{1}{12}}$

関数  $T_f(f, f_c)$  と，関数  $F_f(f)$  との整合度  $C_f(f_c)$  を以下の計算式により求め， $C_f(f_c)$  の最大値を与える  $f_c = f_{c0}$  をその歌唱でのドの基準周波数と推定する． $f_c$  の変化幅はサンプリング周期の逆数とする．

$$C_f(f_{c0}) = \max C_f(f_c) \quad (9)$$

$$\text{ただし， } C_f(f_c) = \int T_f(f, f_c) \cdot F_f(f) df \quad (10)$$

また，ドから半音で  $k$  個上の階名の周波数を，

$$f_k = f_{c0} \times r^k \quad (k = 0, 1, 2, \dots, 11) \quad (11)$$

とする．ただし，

$$\begin{cases} k=0: \text{ド} \\ k=2: \text{レ} \\ k=4: \text{ミ} \\ k=5: \text{ファ} \\ k=7: \text{ソ} \\ k=9: \text{ラ} \\ k=11: \text{シ} \end{cases} \quad (12)$$

である．

ある 1 つの音と判定された時間区間内での基本周波数の平均値を  $f_0$  とすると，

$$f_k \times r^{-\frac{1}{2}} < f_0 \leq f_k \times r^{\frac{1}{2}} \quad (13)$$

を満たすときの  $k$  を，基音ドから半音単位で数えた階名音 ID であると推定する．

### 3.5 出力

結果を音符情報として出力する．本システムでは，MIDI ファイルとして出力する．音高は中央 C を 64 として，そこから半音毎に  $\pm 1$  を加えた数で表示される．現状では長調なら八長調，短調ならイ短調で出力されるが，正規の調として出力する際には固定値を加えるだけで処理できる．

## 4. 評価実験

### 4.1 評価実験の概要

アルゴリズムの動作を確認するために評価実験を行う．以下の 3 つの入力音に対する動作の確認を行

う．実験に使用する曲は，簡単な 7 曲の童謡(最初の 4 小節)である．

( )MIDI 出力音

MIDI 打込みの出力音(音色はピアノ音)に対する動作確認を行う．これは，人が正確な音高で歌うことができないため，音価・音高ともに正確な入力音に対するシステムの動作を確認することが目的である．

( )調弦をずらせたギター

本システムでは，利用者にとっての音階の基準周波数を調べて，それをもとに階名の推定を行っている．この処理が正しく動作しているかどうかを確認するために，調弦を全体にずらせたギターによる演奏に対する動作確認を行う．ここでは，A4=430Hz(正規の周波数より約 40 セント下方へずれている)として調弦したギターを用いて人間が演奏し，それを自動採譜することによって，本システムの動作確認を行う．

( )人による歌唱

人の歌唱に対するシステムの動作を確認する．被験者は 5 名で，年齢は 20 歳前後，歌唱に対する特別なトレーニングなどを受けているような人は含んでいない．録音は比較的静かな部屋で行う．

ハミングで歌うと，音量が小さくなってしまう場合は，基本周波数抽出などの際にその精度が落ちてしまうことを考慮し，「アー」や「ラー」など，母音を含む有声音を何か発声してもらうように，被験者に指示する．

### 4.2 評価基準

評価実験における評価基準を述べる．

人が思い描いている曲は，正規の譜面とは違っている場合がある．例えば，譜面上では四分音符となっている所を，八分音符 2 つであると間違えて記憶している場合などがある．あるいは，八分音符のあとに八分休符が付く所を四分音符で歌うこともある．このような状況が存在するため，正規の譜面と出力結果を比較するのではなく，入力音声と出力結果とを聞き比べ，想定した譜面と比べて正しく認識しているかどうかを評価することとする．

2.1 節で述べたように，本システムでは，休符を認識しないので，休符を除いた音符の総数を用いて，以下の式により認識率を計算する．

$$\text{音価の認識率} = 1 - \frac{\text{音価の誤認識の数}}{\text{想定譜面における音符の総数}}$$

$$\text{階名の認識率} = 1 - \frac{\text{階名の誤認識の数}}{\text{想定譜面における音符の総数}}$$

### 4.3 評価実験の結果

本システムの認識率と、市販ソフト「鼻歌ミュージシャン2」の認識率の比較を以下に示す。

表1 認識率の比較

		MIDI 出力音	ギター	歌唱
本システム	音価の認識率	100%	89%	95%
	階名の認識率	100%	98%	74%
市販ソフト	音価の認識率	53%	59%	69%
	階名の認識率	88%	0%	52%

市販ソフトにおけるギター音の階名の認識については、多分、調弦をA4=430Hzとしたからであろうが、全ての音が同じ音高(A3 前後)と認識され、正確な認識は1つも得られなかった。

人による歌唱の階名の認識率については、人が正確な音高で歌えていない場合が含まれるため、正確な認識率とは言えない。そこで、被験者一人一人について、音階の基準周波数からどの程度ずれて歌っているかを定量的に示すため、正規の音階の基準周波数からの標準偏差(セント値で表現)を求める。また、標準偏差に対するシステムの認識率を下図に示す。

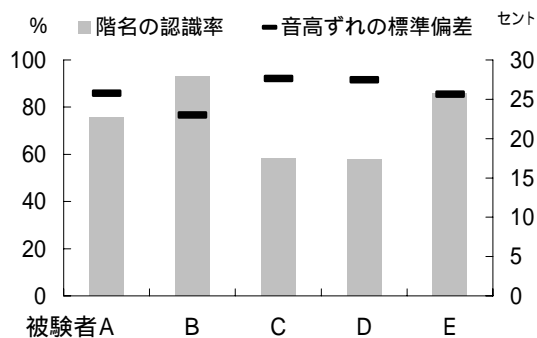


図3 音高のずれとシステムの認識率の関係

図3より、音高のずれの標準偏差が23セント程度の場合、95%前後の認識率を得ることができる。逆に、28セント程度の場合、認識率は60%弱にまで落ち込んでしまう。

システムの誤認識として多く見られるパターンとして、以下の4つが挙げられる。

- ( ) 正規の譜面には存在しない音符が挿入されている。
- ( ) あるべき音符が脱落している。
- ( ) 同じ階名の隣り合う複数の音符を1つの音符と認識してしまう。
- ( ) 1つの音符を複数の音符として認識してしまう。

### 5. まとめと今後の課題

前節の評価実験の結果より、MIDIによる演奏、また、正確な音高で歌えている歌唱については、精度の良い推定が行われていることが分かる。今後の課題としては、音高が正確ではない、人の歌唱に対して、どのような補正を行えば、歌った本人が意図する曲に近い出力を行えるようになるかを検討しなければならない。以下に、今後の具体的な課題を述べる。

( ) 音階の基準周波数の決定において、現在のシステムでは、入力された歌全体に対して、1つの音階の基準周波数を決定している。しかし、人が歌う場合、その人にとっての音階の基準周波数は曲全体で一定であるとは限らず、徐々にずれてしまう場合もあるかもしれない。そこで、曲全体で音階の基準周波数を決定するのではなく、途中で補正しながら、音階の基準周波数を変更していく方法を検討する。

( ) 基本周波数の抽出にケプストラムを採用しているが、音声信号に対して、精度の良い基本周波数抽出を行えるとされている、SIFT, STRAIGHTの採用を検討する。

( ) 今回行った評価実験では、間違えた音高で歌った音は、システムの誤認識とみなして認識率を計算している。システムとしての正確な認識率を得るために、これらを歌う人の責任とするように評価方法を検討しなければならない。ただし、その場合は、正しく歌えなかったことの保障をマニュアルで修正できるような機構が必要である。また、今回の実験は、比較的静かな部屋で取ったが、雑音の多い環境下では、どの程度の認識率を得ることができるのかを把握するための追加実験を行いたいと考えている。

本研究は、同志社大学学術フロンティア事業の援助を受けた。

### 参考文献

- [1] 三浦雅展, 柳田益造他: “ポップス系の旋律に対する和声付与システムAMOR”, 日本音響学会 音楽音響研究会 MA2003-7(2003)。
- [2] 片寄晴弘, 井口征士: “知的採譜システム”, 人工知能学会誌, Vol.5, No.1, pp.59-66(1990)。
- [3] 小林政義, 柳田益造: “時間連動性を考慮した多重調波構造モデルによる多重奏音の自動採譜に関する検討”, 日本音響学会 音楽音響研究会 MA2003-41(2004)。
- [4] Shirado, T. and Yanagida, M.: “Extraction of Fundamental frequencies from the sound of a violin duet”, Proc. of the Joint meeting of ASA/ASJ, pp.291-396(1996)
- [5] 鼻歌ミュージシャン2, (株)メディア・ナビゲーション