

歌唱支援システム構築のための歌声の分析と評価

片岡 靖景 伊東一典 池田操* 中澤達夫** 米沢義道 今関義弘 橋本昌己

信州大学工学部情報工学科 *上越教育大学 **長野工業高等専門学校

vkataoka@cs.shinshu-u.ac.jp itoh@cs.shinshu-u.ac.jp

pierre@ei.nagano-nct.ac.jp yonezawa@cs.shinshu-u.ac.jp

imazeki@cs.shinshu-u.ac.jp hoshimoto@cs.shinshu-u.ac.jp

近年、カラオケの普及により、一般の人が人前で歌う機会が増加しており、上手に歌いたいと練習に励む姿を見かける。また声楽を学ぶ人も自分の習熟度を客観的に評価したいと考えている。そこで本研究では、声楽経験の異なる被験者の歌声を分析して、音高と音長の精度、音量に着目した強弱記号の表現など、種々の歌唱情報を抽出して、習熟度の程度を伝える歌唱支援について検討したので報告する。

Analysis and evaluation of singing voice to produce a support system for singing

Yasuaki KATAOKA Kazunori ITOH Misao IKEDA* Tatsuo NAKAZAWA**
Yoshimichi YONEZAWA Yoshihiro IMAZEKI Masami HASHIMOTO

Department of Information Engineering, Shinshu University
500 Wakasato, Nagano, 380-8553 Japan

*Joetsu University of Education
1 Yamayashikimachi, Joetsu, 943-8512 Japan

**Nagano National College of Technology
716 Tokuma, Nagano, 381-8550 Japan

In recent years, many average people have begun studying how to sing because, with the surge in popularity of Karaoke, they have many chances to sing in front of an audience. On the other hand, people who professionally study vocal music want to receive an objective evaluation of their singing. The instrumental assessment of singing development requires musically relevant stimuli which give accurate information about musical pitch, length of tone, loudness control and other vocal skills. In this paper, we propose a support system for singing evaluation based on the vocal analysis of subjects whose musical experience is markedly different.

1 はじめに

近年、音楽の分野においてコンピュータが広く利用され、様々な音楽情報処理の研究が行なわれている。特に歌唱に関する研究では、歌唱の自動採譜¹⁾、歌声に対する自動伴奏システム^{2),3)}、視覚フィードバックによるトレーニングシステム^{4),5)}などがある。

また、カラオケの普及により一般の人が人前で歌う機会が増加している。しかし、歌の苦手な人には大変な問題である。一方、声楽を専門に学ぶ人にとってどのように声を出すかは大きな問題で、歌唱練習を支援したり、習熟度を客観的に伝えるシステムの構築が望まれている。歌声(歌い声⁶⁾、歌の声⁷⁾)の音響学的研究は多く行なわれており、音声分析による発声の評価も試みられているが^{8),9)}、歌声の分析結果をどのように評価するかは今後の課題である。

ところで、歌唱障害についてはトレーニングによって障害が改善されること、特に音高情報のフィードバックの有効性が知られている^{10),11)}。しかし、声楽を専門に学ぶ人への支援や合唱の歌唱練習支援のためには従来のピッチ抽出や音長の抽出に加えて音量の抽出や他の歌唱技術に関連する情報の抽出が必要である。

そこで本研究では、習熟度の程度を客観的に伝える歌唱支援システム構築のための基礎段階として、どのような歌唱情報の抽出が必要かを検討するために、ピッチと音長の抽出に新たに音量抽出をシステムに加えて声楽経験の異なる被験者の歌声を分析したので、音量分析を中心にこれらの結果について報告する。

2 歌唱支援システムの概要

図1に歌唱支援システムの概略図を示す。先生(声楽経験者)および生徒(声楽未経験者)の歌声はコンデンサマイクホン(STANDARD CMP-115)、音声入出力ボード(CANOPUS Sound Master V)を通して、サンプリング周波数8kHz、量子化8bitでパーソナルコンピュータ(NEC PC-9821Xs)に入力されて、歌唱情報の計測処理が行なわれる。歌唱情報計測処

理部ではピッチ抽出、音長抽出、音量抽出、スペクトル分析などの歌声の特徴抽出が可能である。

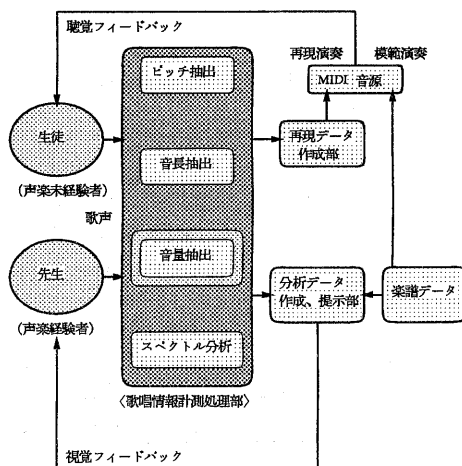


図 1: 歌唱支援システムの概略図

ピッチ抽出は種々の検討結果から、分析長は64ms(512ポイント)で時間窓には折り返し窓を使用し、ピッチ周波数が500Hz以下ではケプストラム法を、500Hz以上では複素スペクトル内挿法を用いることにした。この操作を繰り返すことにより、音高精度の時間変化が分かる。

図2に新たに導入した音量抽出の方法を示す。64ms単位で2乗の平均値を求め、測定区間のパワーの最大値を100として正規化した。

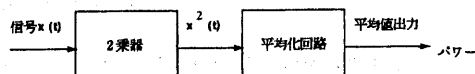
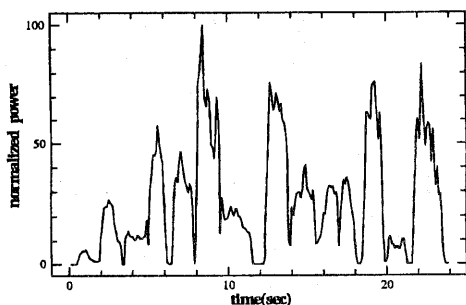


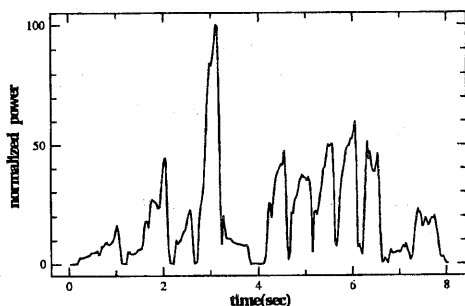
図 2: 音量抽出の方法

ところで、歌唱の分析では様々な速さに的確に対応することが必要で、音量の抽出に当たっても、曲のテンポに応じた処理が可能ないように、テンポ指定を設けた。テンポ60(4分音符)を

基準として、この時のパワーは $T=64\text{ms}$ 毎に抽出されるが、その前後のテンポでは T に反比例するようにした。図 3 (a)、(b) は遅いテンポ ($\downarrow=40$) と速いテンポ ($\downarrow=120$) できらきら星を歌った場合について音量抽出の結果を示したものである。両者ともにデータサイズはほぼ同じで、歌唱の速さに応じた音量の変化の様子が分かる。



(a) テンポ 40



(b) テンポ 120

図 3: きらきら星の音量変化

歌唱情報計測処理部の分析結果は分析データ作成、提示部を経て先生および生徒に視覚フィードバックされる。また、必要に応じて楽譜データとの対比が可能で、音高精度、音長精度も算出できる。一方、聴覚フィードバックについては楽譜データをもとに MIDI 音源 (Roland CM-500, YAMAHA MU100R) による模範演奏および再現データ作成部で作成された先生およ

び生徒のピッチと音長の情報による再現演奏が可能である。再現データ作成部では 64ms 毎に算出されるピッチデータを順次読み込み、75セント以上の差 ($3/8$ 音のずれ) が出るまでのデータを一組として、この中から最も高い頻度で出現するピッチ周波数を選び代表ピッチとする方式¹¹⁾を採用している。

3 強弱記号についての基礎実験

3.1 はじめに

はじめに音楽の強弱記号として、*f*、*mf*、*mp*、*p*、*cresc.*、*decresc.* を取り上げ、声楽経験の異なる被験者の歌声を音量抽出とピッチ抽出を併用して分析して、習熟度の程度について検討する。

3.2 実験方法

被験者は声楽経験者 8 名 (男性 4 名、女性 4 名) および声楽未経験者 8 名 (男性 8 名) で、各被験者についての音楽歴を簡単に示すと次の通りである。

経験者 A (女): 音楽 (声楽) 科の先生

経験者 B (女): 音楽 (声楽) 科の学生
合唱団にも所属 (3 年間)

経験者 C (男): 合唱団に所属 (4 年間)
専門に声楽は学んでいない

経験者 D (男): 合唱団に所属 (10 年間)
専門に声楽は学んでいない

経験者 E (女): 音楽 (声楽) 科の学生

未経験者 F~I (男): 声楽に関する経歴はない

実験項目は以下に示す通りである。

<実験 1 >

発声は /a/ でこれらの被験者の発声しやすいピッチ周波数として男性は 196Hz (音階でソ) で、女性は 660Hz (音階でミ) である。またテンポ 80 で (*f* → *mf* → *mp* → *p*) の順に 2 拍毎に発声と休符が繰り返される。同様に (*p* → *mp* → *mf* → *f*) と逆順の発声も合わせて行なう。

<実験 2 >

cresc. → decresc. の実験では8拍の中で前半4拍がクレシェンド、後半4拍がデクレシェンドで発声は/a/で、ピッチ周波数は実験1と同じである。

<実験3>

実際に強弱記号 cresc. と decresc. のある曲として「春の小川」(文部省唱歌、テンポ104)を選曲し、その一部分を歌い、歌唱時の音量分析を試みる。この時、cresc. と decresc. は以下のように曲に付いている。

(cresc.) (decresc.)
 は一のおがわは | さらさらいくよ

被験者はいずれも、この曲を過去に歌った経験があるが、久しぶりに歌うということで、音、リズム、楽譜の強弱記号 cresc. と decresc. 等の確認を自分が納得するまで行なう。

3.3 実験1の結果と考察

ここでは強弱記号の発声順序を変えた2つの方法を試みたが、f → mf → mp → p という音量的に下降していく場合に特に声楽経験の差異が見られたので、それらの被験者の結果を中心に考察する。図4は経験者から1人(C)、未経験者から2人(F、I)を選び、実際の音量データを重ねて示したものである。図5はその時のピッチ変化を示している。図4より経験者Cはそれぞれの発声においてほぼ一定の音量を保ちつつ、強弱記号の発声順に対応して音量が下がっているのが分かる。これに対し未経験者F、Iは、それぞれの発声において一定の音量ではなく、初めが強く次第に下がる傾向が見られる。また未経験者Iは経験者Cと同様に、各強弱記号に対応して音量が下がっていくが、未経験者Fの場合は、mfとmpの音量差がないことが分かる。

ピッチに関しては図5より、経験者Cは、実験条件に近い190 Hz付近でほぼ安定している。未経験者Fはピッチはほぼ安定しているものの発声時間が短く、テンポ80より2拍の発声時間がおよそ1.5 secぐらいと考えられる。未経験者Iはmfからピッチが下がり、p

の発声時間が短くなっており、強弱記号による影響が見られる。

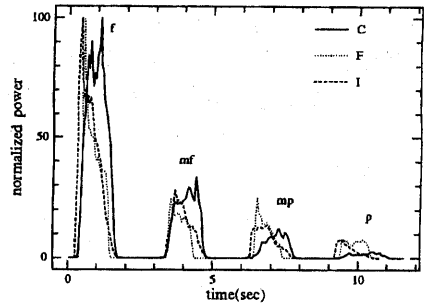


図4: f → mf → mp → p 時の音量変化

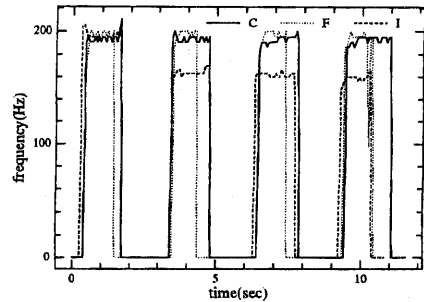


図5: f → mf → mp → p 時のピッチ変化

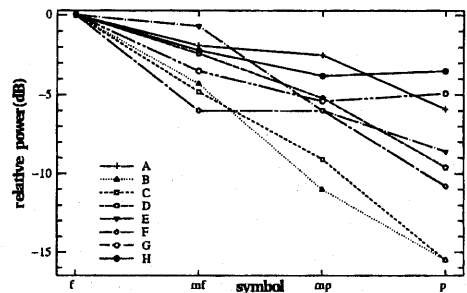


図6: 強弱記号による音量のピークの変化

図6は経験者5人(A,B,C,D,E)、未経験者3人(F,G,H)について各強弱記号での音量のピークを検出し、縦軸をデシベル(dB)表示にしてその変化の様子を示したものである。ここで、声楽経験者と合唱経験者では比率は違うものの各強弱記号に対応して音量が下がっているのが分かる。比率が違うことについて、声楽経験者A、Eは独唱を専門とするのでpでもある程度の音量が出ているが、合唱経験者B、Cは各パート全体のハーモニーで音量を考慮するためか、mp、pではかなり音量が小さくなっている。一方、未経験者F~Hは各強弱記号に対応した結果が見られず、未経験者G、Hのようにmp→pにおいて音量が上がる場合や、未経験者Fのようにmf→mpで音量がほとんど変化しない場合が見られる。

3.4 実験2の結果と考察

次にcresc. → decresc.の結果を示す。図7は実験結果から代表的な特性を示す経験者2人(A、C)、未経験者2人(G、H)を選び、実際の音量データを重ねて示したものである。図8はその時の男性の被験者(C、G、H)のピッチ変化を示している。なおcresc. → decresc.の音量変化を模擬した波形では、ほぼ2次関数の上昇と下降の特性が得られることを確認している。

声楽経験の差異は、主に図7に示すように音量の変化特性パターンの差として表れている。経験者A、Cは、cresc. → decresc.に伴い模擬パターンに近い形が見られる。これに対して未経験者G、Hは、最初の部分での凹凸が顕著で、発声した瞬間に一時的に音量が大きくなり一旦減少した後、cresc. → decresc.の特性を示している。この原因として未経験者は、発声技術が未熟なため、なめらかな音量の上昇、下降の制御が困難と考えられる。また未経験者では特性に個人差が多く見られる。

ピッチに関しては図8より、経験者Cは、実験条件に近い190 Hz付近ではほぼ安定していることが分かる。これに対して未経験者Hはピッチが明らかに高く、未経験者Gと共に、発

声時間も短いことが分かる。

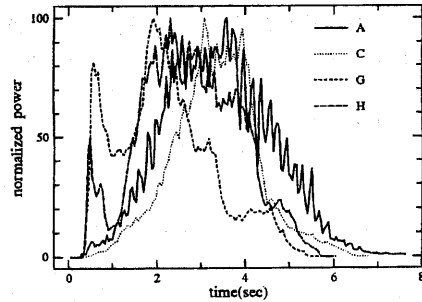


図7: cresc. → decresc. の音量変化

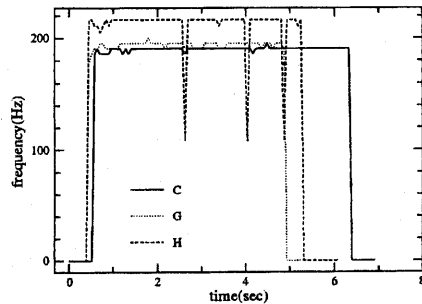
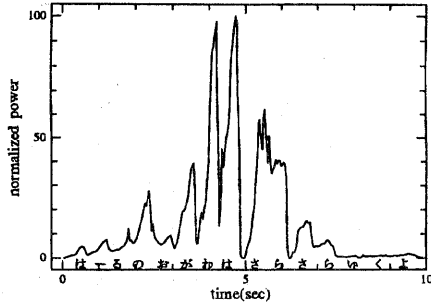


図8: cresc. → decresc. のピッチ変化

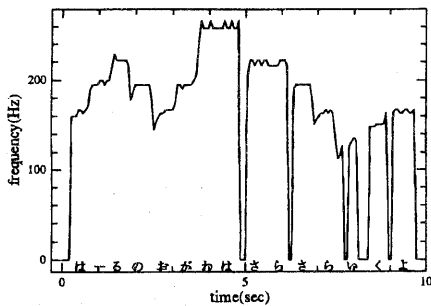
3.5 実験3の結果と考察

次に「春の小川」についての歌唱時の結果を示す。実験結果から経験者1人(C)、未経験者1人(G)を選び、各々の音量データとピッチデータを図9(a)、(b)および図10(a)、(b)に示す。ピッチに関しては、声楽経験の差がほとんど見られないが、音量については明らかに特性に違いが見られ、その傾向は前半のcresc.の部分で顕著である。両被験者とも音量変化にcresc.とdecresc.を意識した結果が表れているが、図9(a)の経験者Cの場合には、全体的にcresc. → decresc.に従った比較的なめらかな音量の変化が見られる。これに対して図10(a)の未経験者Gの場合には、図10(b)に示すピッチの変化の影響を受けてピッチの高い部

分で、音量が大きくなる傾向が見られる。以上の結果から歌唱時においても、先の *cresc.* → *decresc.* の実験結果同様に音量分析が声楽経験の差を見るために有効であることが分かる。

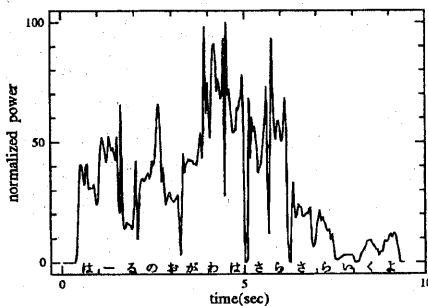


(a) 音量

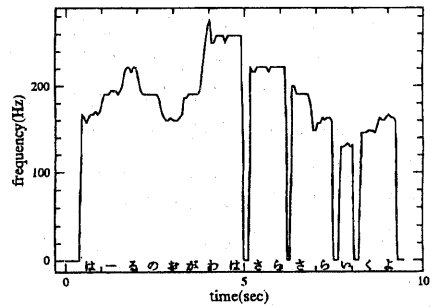


(b) ピッチ

図 9: 経験者 C の歌唱分析



(a) 音量



(b) ピッチ

図 10: 未経験者 G の歌唱分析

4 奏法についての基礎実験

4.1 はじめに

次に、基礎的な奏法記号を伴う歌唱について、前述の実験と同様な分析方法で検討することにした。ここでは奏法標語のスタカートとテヌートを取り上げ、実験を行なう。

4.2 実験方法

被験者は強弱記号の実験から男性について代表的な声楽経験者と未経験者各々 1 人で、実験項目は以下に示す通りである。

<実験 1>

ピッチ周波数、発声、テンポは強弱記号の実験条件と同じである。スタカート→普通の順にそれぞれ 3 回ずつ発声し、1 拍毎に発声と休符が繰り返される。また未経験者にはスタカートは、歯切れよく音と音との間を切って発声するように指示してある。

<実験 2>

実験 1 と同様な条件で、スタカートをテヌートに変えて実験を行なう。また未経験者にはテヌートは、その音の長さを十分に保って、丁寧に発声するように指示してある。

4.3 実験1の結果と考察

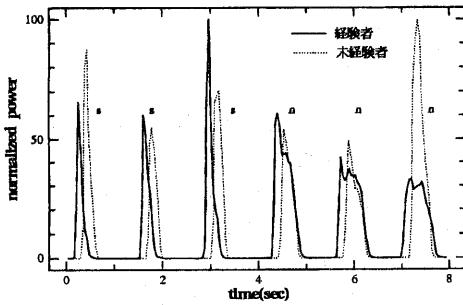


図 11: s → n の音量変化

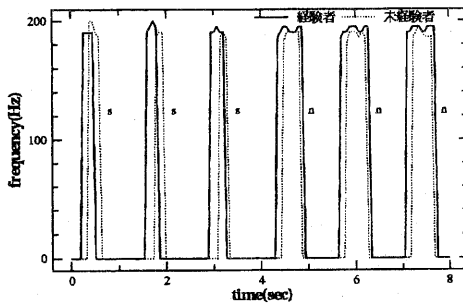


図 12: s → n のピッチ変化

図 11 は経験者と未経験者の音量データを重ねて示したものである。s はスタカートの発声を、n は普通の発声を示している。図 12 はその時のピッチの変化を示している。図 11 より経験者については、音量変化にスタカートの基本的な特性として考えられる発声時間が短く、鋭い形状が見られる。これに対して未経験者も、スタカートの特有の形状が見られるが、普通の発声部分でもそれに類似する形状が見られる。このことは前半のスタカートの発声が後半の普通の発声に影響を与えているものと考えられる。

図 12 に示すピッチについては両者とも実験条件に近い 190Hz 付近で安定して発声していることが分かる。ここで、発声時間に注目すると、経験者はスタカートの発声時間と普通の発

声時間の比率が、0.50 であるのに対して、未経験者はその比率が 0.57 であり、未経験者の普通の発声時間がスタカートの影響を受けて短くなっている。この実験はテンポ 80 で行なっているので、普通の発声時間が 0.7~0.8sec となるのが標準で、経験者は 0.68sec であるのに対して未経験者の場合にはその発声時間の平均は 0.49sec であった。

4.4 実験2の結果と考察

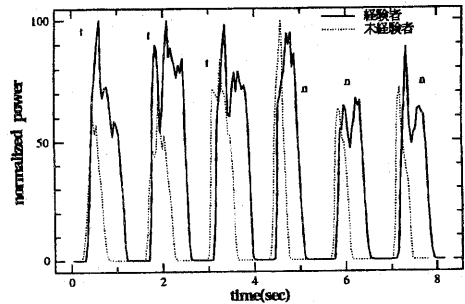


図 13: t → n の音量変化

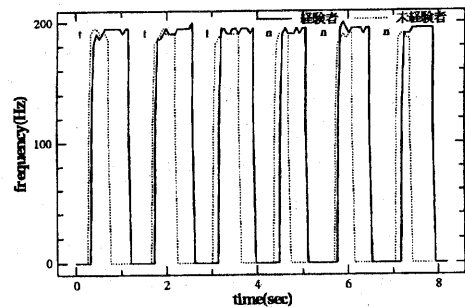


図 14: t → n のピッチ変化

図 13 はテヌート→普通の場合の結果を経験者と未経験者について音量データを重ねて示したものである。図 14 はその時のピッチ変化を示している。ここで t はテヌートの発声を、n は普通の発声を示している。図 13 より経験者はテヌートの基本的な特性として考えられる発声時間が長く、その間、ほぼ均等なパワーを保

つ形状が見られる。これに対して未経験者はテヌートと普通での発声の差異がほとんど見られない。図14に示すピッチについては両者とも実験条件に近い190 Hz付近でほぼ安定して発声しているが、発声時間に注目してみると、経験者はテヌートの発声時間と普通の発声時間の比率が1.19であるのに対して、未経験者はその比率が1.41である。ここではテンポ80で実験を行なっているので、普通の発声時間の標準は0.7~0.8secとなる。経験者のそれが0.73secにあるのに対して、未経験者は0.41secであり、未経験者がテヌートから普通の発声へと移る時に普通の発声時間を短く発声していることが分かる。

5 まとめ

習熟度の程度を客観的に伝える歌唱支援システムを構築するために必要な歌唱情報を抽出することを目的として、ピッチ抽出と音長抽出に新たに音量抽出を加えた。声楽経験の異なる被験者に強弱記号や奏法記号に注意して歌唱してもらい、それらの歌声を分析して、習熟度の程度を判断する要因について種々検討した。得られた知見をまとめると次の通りである。

強弱記号の発声については、声楽経験者の方が未経験者に比べて優れているので各強弱記号に対応した適切な音量変化が見られた。またピッチの変化においても声楽経験者の場合は未経験者に比べてピッチ精度、音長精度ともに優れていることが分かった。また未経験者については特性に個人差が見られることからある程度、習熟度の経過も推定できるのではないかと考えられる。単音の発声と同様に簡単な歌唱についても *cresc.* と *delesc.* の特性に声楽経験の差が確認できた。

スタカートとテヌートの奏法の違いによる発声においても声楽経験者と未経験者の特性には差異が確認でき、これらの場合についても習熟度の程度の判断の可能性が示された。

今後、さらにこれらの歌唱情報の被験者への有効なフィードバック法についても検討を重ね

たい。

実験に御協力頂いた掘 和文君に感謝します。

参考文献

- 1) 新原, 今井, 井口:歌唱の自動採譜, 計測自動制御学会論文集, 20,10,pp.940-945(1984).
- 2) 井上, 橋本, 大照:適応型歌声自動伴奏システム, 情報処理学会論文誌,37,1,pp.31-38(1996).
- 3) 東, 橋本:音声認識とピッチ抽出を併用した歌声の自動伴奏, 情報処理学会研究報告,97-MUS-22,pp.1-5(1997).
- 4) D.M.Howard and G.F.Welch:Visual Displays for the Assessment of Vocal Pitch Matching Development, Applied Acoustics,39,pp.235-252 (1993).
- 5) 片寄, 金森, 平井, 村尾, 井口:MAX を利用した Voice Shooting Game, 情報処理学会音楽情報科学研究会資料 11-7,pp.45-48(1995).
- 6) 安藤 由典:新版 楽器の音響学, 第11章 歌い声, pp.235-246, 音楽之友社(1996).
- 7) 中山, 小林:歌の声—声質の魅力と問題点—, 日本音響学会誌,52,5,pp.383-388 (1996).
- 8) 吉田 泰昌:音声分析による声楽発声の物理音響学的評価, 日本声楽発声学会誌,23, pp.5-13(1995).
- 9) 池田 操:音響分析による歌曲「赤とんぼ」の歌唱評価, 上越教育大学研究紀要,17 ,1,pp.395-407(1997).
- 10) 米沢, 伊東, 平野:音痴の音高感覚に関する検討, 電子情報通信学会論文誌 A, J71-A,8,pp.1532-1538(1988).
- 11) 伝田, 横山, 伊東, 橋本, 今関, 米沢:歌唱情報の音高フィードバックに関する検討, 平成8年度電子情報通信学会信越支部大会, V29,pp.57-58(1996).