

アマチュア歌唱エンタテインメントのための 熱唱度評価システム SEES

大道 竜之介^{†1} 伊藤 仁^{†1,*1}
牧野 正三^{†1,*2} 伊藤 彰 則^{†1}

本研究の目標は、歌唱の熱唱度を自動的に評価するシステムを開発することである。カラオケの採点システムに代表されるように、歌唱音声の自動評価によってアマチュア歌唱エンタテインメントに付加価値をもたらそうとする研究が盛んに行われてきた。しかしほとんどの研究において、その評価しようとする側面は歌唱のうまさである。本研究では歌唱のうまさではなく熱唱度に着目し、その知覚に関する基本的な特性を明らかにするとともに、自動評価システムの開発を目指す。またこれを用いた新しい形態のアマチュア歌唱エンタテインメントを提案する。

The SEES; Singing Enthusiasm Evaluation System for Amateur Singing Entertainment

RYUNOSUKE DAIDO,^{†1} MASASHI ITO,^{†1,*1}
SHOZO MAKINO^{†1,*2} and AKINORI ITO^{†1}

The goal of our research is to develop a system for evaluating singing enthusiasm. As evaluation systems for karaoke represent, many researchers have worked on automatic evaluation methods of singing voice to make additional value on amateur singing entertainment. However most of the researches try to evaluate only singing skill. In our research, the point of interest is not singing skill but singing enthusiasm. We describe in this paper our attempt to develop an automatic evaluation system of singing enthusiasm through analyses of principles on human perception of it. Moreover we propose a new style of amateur singing entertainment with our system.

1. はじめに

音楽には極めて多様な形態があるが、その中でも歌唱は、人間に本来的に備わった発声器官を用いて誰でも楽しむことができる点で特別な形態である。声楽家やポピュラー歌手といったプロフェッショナルによる演奏を聴いて楽しむことが文化を形成する一方で、アマチュアすなわちプロフェッショナルでない一般人が、自己もしくは近いもの同士の余暇活動として歌を歌うことを楽しむ文化が存在する。このような文化を本稿ではアマチュア歌唱エンタテインメントと称し着目する。

アマチュア歌唱エンタテインメントの代表的な例はカラオケである。カラオケは1970年代に日本で生まれた。『レジャー白書 2011』¹⁵⁾では国民の主な余暇活動として21の娯楽を挙げ参加人口等の統計を掲載している。「カラオケ」の参加人口はその中で第2位の4,680万人と推計され、第1位である「外食」の6,040万人に次いで大きい。なおかつカラオケは今や世界各国で主要な余暇活動のひとつに数えられる⁷⁾。カラオケを代表とするアマチュア歌唱エンタテインメントにさらなる価値を付加するべく多様な研究が行われており^{1),14)}、中でも歌唱の自動評価に関する研究が多数ある^{5),6),8),12)}。このような研究のほとんどは、歌唱の「うまさ」を評価しようとするものであり、多くの場合、歌唱音声から音高の正確さを含むいくつかの音響的特徴を抽出することを基本原理とする。しかしアマチュア歌唱エンタテインメントにおいて歌唱音声の評価すべき側面は必ずしもうまさに限られないと我々は考え、新しい評価側面として「熱唱度」を提案した¹³⁾。

以下に本稿の構成について述べる。まず第2節では我々の研究における熱唱度の概念を、歌声に関する他の概念との比較を含めて説明する。第3節では実験に用いた歌唱音声コーパスについて述べ、人間が歌唱音声の聴取によって熱唱度を知覚する際の要因や、人間同士の評価の共通性について議論する。第4節では熱唱度の知覚に関係すると考えられる音響的特徴を挙げ、その定量化方法について述べる。第5節では我々が開発した熱唱度評価システム *SEES* (*Singing Enthusiasm Evaluation System*) の概要および性能評価について述べる。第6節では *SEES* を実装した iPhone アプリを挙げ、新しい形態のアマチュア歌唱エン

^{†1} 東北大学

Tohoku University

*1 現在、東北工業大学

Presently with Tohoku Institute of Engineering

*2 現在、東北文化学園大学

Presently with Tohoku Bunka Gakuen University

タインメントとしての *SEES* の応用を提案する。最後に第 7 節で本稿のまとめを述べる。

2. 熱唱度

2.1 熱唱度とは

本研究で我々が述べる「熱唱度」とは歌唱に対する態度のひとつとして、いかに一生懸命に歌唱しているかを表す度合いである。歌唱音声の分析に関しては「歌唱力」や「感情表現」といった側面についての研究がこれまでに行われてきた。以下にこれらの概念と熱唱度とを比較しその違いについて述べる。

第 1 節で述べた通り、歌唱力と音響的特徴との関係について分析する研究や、歌唱力を自動的に評価しようとする試みが近年複数報告されている。これらは歌唱のうまさに着目した研究である。歌唱のうまに関する明確な定義は存在しないが、うまい歌唱の必要条件として、少なくとも聴く人に不快感をもたらすような歌唱でないことが挙げられるだろう。歌のうまさと熱唱度とはこの点で明確に区別することができる。ときにマスターセッションと評されるような自分本位でへたな歌唱であっても、一生懸命な歌唱すなわち熱唱であり得る。

歌唱における感情表現と音響的特徴の関係性についても、いくつかの研究が行われている^{2),4),9),10)}。Justin らは、音声発話と音楽表現における感情に関して過去に行われた研究をまとめた³⁾ 中で、基本的な感情を分類する言葉は怒り (Anger)・恐れ (Fear)・喜び (Happiness)・悲しみ (Sadness)・愛情 (Love-tenderness) の 5 種類に集約できるとした。これらの言葉にみられるように、感情とは基本的に心理的状态の分類であり、一生懸命さという歌唱の態度に着目した熱唱度はこれと異なる概念である。

2.2 知覚熱唱度と本人熱唱度

本研究では熱唱度の観点として「本人熱唱度」と「知覚熱唱度」の 2 つを導入し、これらを明確に区別する。「本人熱唱度」は歌唱者本人が意図する熱唱度、「知覚熱唱度」は歌唱者以外の人々が歌唱音声を聞いた際に知覚する熱唱度とする。熱唱度評価システムにおいて我々が客観評価しようとするのは知覚熱唱度であり、人が聞いて知覚する熱唱度に近い値を評価するシステムを目指す。本人熱唱度と知覚熱唱度との関係性については第 2 節で分析する。

3. 歌唱音声コーパスの作成と分析

本研究では 3 セットの歌唱音声コーパスを作成し、それぞれセット A, セット B, セット C と称して用いた。本節ではその作成手順と分析結果について述べる。

3.1 歌唱音声コーパスの作成

3.1.1 歌唱収録

34 人のアマチュア歌唱者が、共通曲としてポピュラーソング『いとしのエリー』1 番を歌唱した。またこのうち 11 人の歌唱者は、さらに自由曲として各自の得意なポピュラーソングの 1 番を歌唱した。幅広い熱唱度の歌唱音声を得るために、共通曲では「熱唱」および「普通」に各 1 回ずつ、また自由曲では各 2 回ずつ歌唱するよう歌唱者に教示し、この 2 水準を本人熱唱度として各音声にラベル付けた。共通曲『いとしのエリー』は下記の要件をみたくように選曲したものであり、自由曲の選曲においても同様の要件を被験者に課した。

- 歌唱者が楽曲のメロディを知っていること
 - 極端な高音や低音がないなど、歌唱の難易度が高すぎないこと
 - 楽曲のムードに左右されすぎずに「熱唱」と「普通」の歌い分けが可能と思われること
- 共通曲のキーは C-Maj. で一定とし、自由曲のキーは歌唱者が歌いやすいキーを自身で選択した。

歌唱収録は防音室内で行い、歌唱者にはヘッドホンで伴奏を提示した。歌唱音声はサンプリング周波数 44.1[kHz]、量子化ビット数 16[bit] で記録した。

3.1.2 フレーズの切り出し

コーパスに用いる音声として、収録した歌唱の中からフレーズを多数切り出した。ここでいうフレーズとは、約 10 秒以内で、歌詞やメロディにまとまりのある歌唱部分の単位を指す。セット A は、同一の歌詞“エリー”および同一のメロディで共通曲に 4 回現れる単一のフレーズを切り出したものである。セット B は共通曲の歌いだし・中盤・サビ・終盤よりそれぞれ 1 フレーズずつ、計 4 フレーズを切り出したものである。セット C は各自由曲より 4 フレーズずつ、歌詞やメロディがなるべく多様になるように切り出したものである。なおセット B の 4 つのフレーズを本稿ではそれぞれ B1 - B4 と称する。

フレーズを切り出した後すべての音声の平均パワーを一定値に正規化した。信号のパワーはマイクと歌唱者との距離や機器の入力レベルに依存するため、本研究では音響特徴量として検討しないためである。

3.1.3 主観評価実験

主観評価実験は表 1 に示す評価語および評価値を用いて以下の手順で行った。

- (1) 練習用刺激を無作為順に提示して評価の練習を行う
- (2) 評価用刺激を無作為順に提示して評価を行う
- (3) (2) を 3 回にわたって繰り返す

表 1 主観評価実験の評価語と評価値

Table 1 Evaluation words and the values for the subjective evaluations

評価語	評価値
熱唱していると感じる	2
どちらでもない	1
熱唱していると感じない	0

表 2 歌唱コーパスの仕様

Table 2 Property of the singing voice corpus

	セット A	セット B	セット C
歌唱者数	34	34	11
楽曲数	1	1	11
フレーズ数	1	4	44
本人熱唱度	2 水準 (「熱唱」 / 「普通」)		
音声数	272	272	176

練習用刺激の個数は 20 個であり、評価用刺激と重複しないように選出した。評価用刺激の個数は 100 個であり、各コーパスに含まれる全音声の評価回数が一定となるように選出した。被験者数はセット A およびセット B では 30 人、セット C では 15 人とした。主観評価実験の後、各刺激に対して全被験者が付与した評価値の平均値を主観評価値としてラベル付けした。

以上の手順により作成した歌唱音声コーパスの仕様を表 2 にまとめる。

3.2 歌唱音声コーパスの分析

3.2.1 主観評価値の分布

各セットごとの主観評価値の分布を図 1 に示す。各セットとも概ね 0.0 - 2.0 の全範囲にわたる主観評価値が付与されていることがわかる。セット B とセット C においては評価値に若干の偏りが見て取れる。これはフレーズを要因とする知覚熱唱度の差異に起因するものと考えられる。詳細は 3.2.3 節で議論する。

3.2.2 被験者同士の相関

各被験者について、その被験者が付与した評価値とそれ以外の被験者が付与した評価値の平均との相関係数を調べた。頻度を正規化した度数分布を図 2 に示す。ほとんどの被験者について 0.7 程度以上の高い相関係数が観測されたことから、人間は歌唱音声の聴取のみによって、おおよそ共通の基準で知覚熱唱度を評価できるものといえる。

3.2.3 知覚熱唱度の要因

セット B の主観評価値について、本人熱唱度・フレーズ・歌唱者およびこれらの 1 次の交互作用を要因とした 3 元配置の分散分析を行った。結果を表 3 に示す。また図 3 に (a) 歌唱者ごと、(b) 本人熱唱度ごと、(c) フレーズごとの主観評価値の分布を示す。これらの箱ひげ図において実線は中央値、箱は第 3 四分位数 (75% 点) および第 1 四分位数 (25% 点)、ひげの両端は以内にある最小および最大値、ひげの外側にある値ははずれ値を示す。図 3(a) では中央値の順序で歌唱者を並べた。

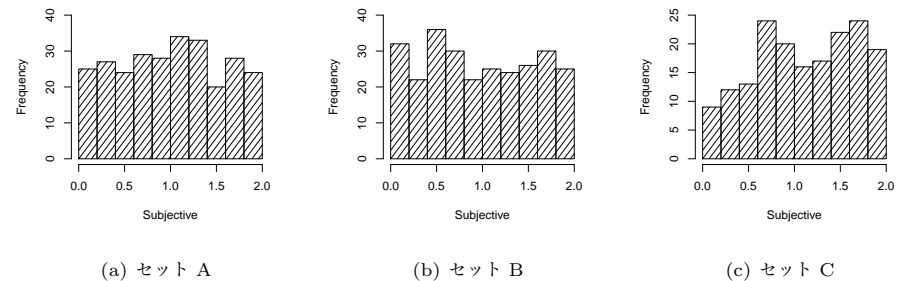


図 1 歌唱音声コーパス各セットにおける主観評価値の分布

Fig. 1 Histograms of the subjective evaluation values for each singing voice corpus set

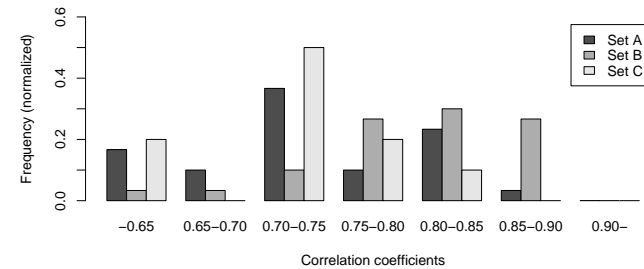


図 2 被験者同士の相関係数

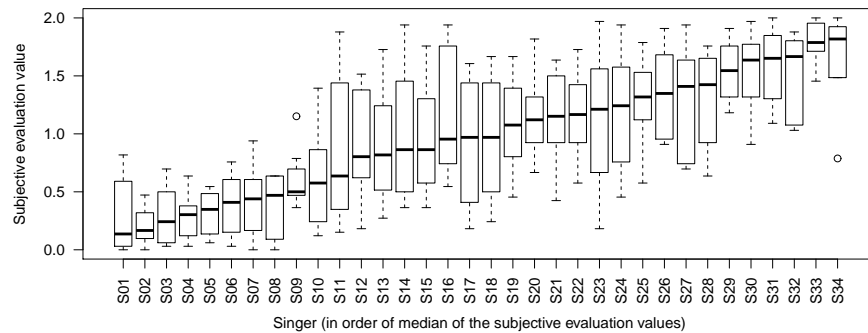
Fig. 2 Correlation coefficients of the subjective evaluation values between the subjects

これらの結果について、以下のように考察する。歌唱者ごとに主観評価値の分布に明らかな差異があることから、声の質や歌い直しなどにおける個人性が知覚熱唱度に影響するものと考えられる。なおかつ本人熱唱度、および歌唱者と本人熱唱度との交互作用による効果が有意であることから、本人熱唱度は知覚熱唱度に影響を与えるといえる。フレーズごとの差異については、図 3(c) よりサビ区間のフレーズ B3 の主観評価値が最も高いことがわかる。この原因としては、フレーズごとの印象が知覚熱唱度に影響すること、また楽曲の構成やフレーズごとの印象が本人熱唱度に影響し間接的に知覚熱唱度に影響することの、どちらか

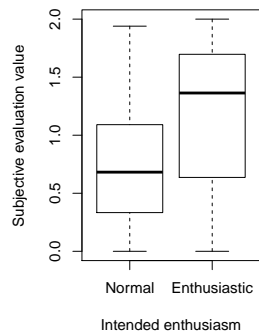
表 3 セット B の主観評価値に関する分散分析表

Table 3 Analysis-of-variance table on the subjective evaluation values of Set B

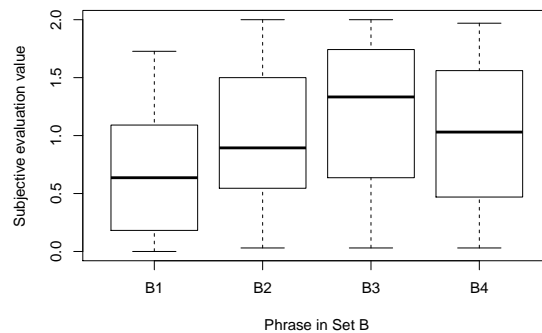
要因	自由度	平方和	平均平方和	F 値	p 値
本人熱唱度	1	13.620	13.6202	346.1223	4.4e-34
フレーズ	3	10.144	3.3815	85.9313	1.9e-27
歌唱者	33	54.566	1.6535	42.0110	7.7e-45
(交互作用) 本人熱唱度×フレーズ	3	0.313	0.1042	2.6478	5.3e-02
(交互作用) フレーズ×歌唱者	99	8.273	0.0836	2.1237	1.1e-04
(交互作用) 歌唱者×本人熱唱度	33	3.736	0.1132	2.8771	2.9e-05
残差	99	3.896	0.0394		



(a) 歌唱者と主観評価値



(b) 本人熱唱度と主観評価値



(c) フレーズと主観評価値

図 3 セット B の歌唱音声における各要因と主観評価値との関係

Fig. 3 Subjective evaluation value for each factors of the singing voice in Set B

しくは両方が考えられる。

4. 音響特徴量

本節では熱唱度自動評価システムに用いる音響特徴量として検討したものを挙げ、その抽出方法および有効性について議論する。なお基本周波数 (F0) は *The Snack Sound Toolkit*¹¹⁾ を用いて 10[ms] 間隔で抽出した。

4.1 原理と抽出方法

4.1.1 A 特性パワー

A 特性パワーは声の大きさに関する特徴量である。主観評価実験の被験者らより、大きな声に対して熱唱度を高く知覚すると報告があった。声の大きさは信号の振幅に影響するが、振幅は歌唱者とマイクとの距離や入力レベルなど他の要因によっても大きく影響されるため、声の大きさを直接に評価する特徴量としては妥当でない。そこで知覚的な声の大きさに影響する周波数成分を評価する特徴量として、この特徴量を考案した。

A 特性パワー特徴量の抽出方法は以下の通りである。まず入力音声信号の平均パワーを一定値に正規化する。その後 A 特性周波数重みをかけて、改めて平均パワーを算出する。

A 特性周波数重みは等ラウドネス曲線の近似特性であり、騒音レベルを測定するための周波数重みとして広く用いられている。本稿の実験において A 特性周波数重みは 1023 点の FIR フィルタにより実装した。

4.1.2 最大 Δ パワー

最大 Δ パワーは、知覚的な声の大きさに関係する特徴のひとつとして、振幅の立ち上がりの急峻さに着目した特徴量である。計算条件はフレーム長 20[ms]、フレームシフト 10[ms]、サイドフレーム数 4 として Δ パワーを計算し、その最大値を抽出した。

4.1.3 ずり上げ・ずり下ろし

ずり上げ・ずり下ろしは図4に示すような、フレーズの先頭および末尾における基本周波数の変動に関する特徴量である。被験者らから、このようなピッチの変動が熱唱度の知覚に影響を与えるとの報告があった。歌唱音声の基本周波数におけるこのような特徴については、歌声の表現手法のひとつとして Mayor ら⁶⁾が挙げているものの、その特性や知覚との関係性については述べられていない。

これらの基本周波数変動を評価する特徴量として、式(1)による二乗平均平方誤差 (RMSE) を抽出した。

$$E_{\text{RMS}}(\tau, T) = \sqrt{\frac{1}{T} \sum_{t=\tau}^{\tau+T-1} (F_{\text{max}}(\tau, T) - F_0(t))^2} \quad (1)$$

$$F_{\text{max}}(\tau, T) = \max_{0 \leq t < T} F_0(\tau + t) \quad (2)$$

ただし $F_0(t)$ は t 番目フレームの基本周波数、 t_s は対象区間の最初の時刻、 T は区間の時間長である。ここで $T=200[\text{ms}]$ とした。

楽譜情報を用いることなく特徴量を抽出するために、ここでは基本周波数が 500[ms] 以上にわたって連続的に検出された区間をフレーズとみなし、全フレーズに対してずり上げおよびずり下ろしに対応する下記の値を計算した。

$$E_{\text{up}} = E_{\text{RMS}}(t_s, T) \quad (3)$$

$$E_{\text{down}} = E_{\text{RMS}}(t_E - T, T) \quad (4)$$

ただし t_s および t_E は、それぞれフレーズの開始時刻と終了時刻である。

4.1.4 ビブラート

多くの被験者が、ビブラートの存在が熱唱度の知覚を向上させると指摘した。中野らが提案する「ビブラートらしさ」⁸⁾を用いてビブラートを検出し、図5および下式(5)-(7)の通りビブラートの深さ V_r 、速さ V_e 、歌唱区間に対する時間割合 V_t を抽出した。

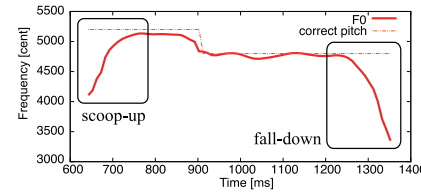


図4 ずり上げ・ずり下ろし
Fig. 4 Scooping-up and falling-down

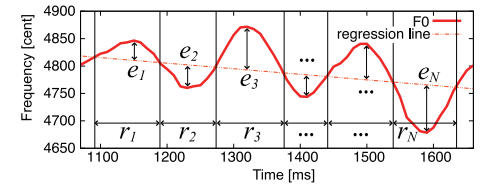


図5 ビブラート特徴量
Fig. 5 Vibrato parameters

$$V_r = \frac{1}{T_v} \sum_{t=1}^{T_v} \frac{1}{2r_t} \quad (5)$$

$$V_e = \frac{1}{T_v} \sum_{t=1}^{T_v} \frac{1}{e_t} \quad (6)$$

$$V_t = \frac{1}{T_{F_0}} \sum_{t=1}^{T_v} r_t \quad (7)$$

ただし T_v はビブラートが検出された区間の時間、 T_{F_0} は基本周波数が検出された全区間の時間である。

4.2 有効性の検討

各音響特徴量と主観評価値との相関係数を、コーパスおよびフレーズごとに調べた結果を表4に示す。A 特性パワー・ずり下ろし・ビブラートに関する特徴量についてはいずれのコーパス (フレーズ) についても正の相関となった。よってこれらの特徴量はフレーズの歌詞やメロディに依存することなく、自動評価に有効に用いることができると考える。一方最大 Δ パワー・ずり上げについては、コーパス (フレーズ) によって負の相関が現れた。よってこれらの特徴量はフレーズによって有効性が一定でなく、自動評価に用いるのに適さないといえる。なおビブラートに関する3つの特徴量については、特徴量間の相関係数が0.7程度以上あり、自動評価に用いるのは深さ V_e のみでよいと考える。

5. 熱唱度評価システム SEES

第4節までの検討に基づき、図6に示す熱唱度評価システム SEES (Singing Enthusiasm Evaluation System) を構築した。本節ではその概要と性能評価について述べる。

表 4 特徴量と主観評価値との相関係数

Table 4 Correlation coefficients between the acoustic parameters and the subjective evaluations

音響特徴量	セット A	セット B	B1	B2	B3	B4	セット C
A 特性パワー	0.47	0.54	0.36	0.50	0.51	0.49	0.62
最大 Δ パワー	0.23	-0.22	0.05	0.13	-0.10	-0.09	-0.16
ずり上げ E_{up}	0.20	0.07	-0.09	0.21	0.14	-0.12	0.10
ずり下ろし E_{down}	0.35	0.36	0.13	0.38	0.29	0.50	0.30
ビブラート深さ V_e	0.37	0.37	0.38	0.27	0.38	0.47	0.32
ビブラート速さ V_r	0.37	0.37	0.39	0.29	0.38	0.47	0.30
ビブラート時間 V_t	0.37	0.30	0.42	0.25	0.30	0.36	0.31

5.1 システム概要

SEES は 3つのパートで構成される。front-end は前処理部であり、低周波ノイズ除去のためのハイパスフィルタ・パワー正規化・基本周波数抽出を行う。基本周波数抽出には 11) を用いた。core は特徴量抽出部であり、4 節に述べた方法に基づき A 特性パワー・ずり下ろし・ビブラート深さの 3つの特徴量を抽出する。back-end は評価値算出部であり、特徴量の重み付き線形和によって最終的な熱唱度評価値を算出する。重み係数 $a_0 - a_3$ はセット A に対する重回帰分析により決定した。

5.2 性能評価

SEES にセット A - セット C を入力した。出力値と主観評価値との関係を図 7 に示す。セット A は back-end で用いる重み係数の決定に用いたため closed な評価、セット B およびセット C はシステムにとって未知の open な評価である。主観評価値との相関係数は相関係数はセット A:0.60, セット B:0.65, セット C:0.64 であった。

第 2 節における被験者同士の主観評価値の相関を調べた図 2 と比較すれば、SEES は人間に迫る評価性能をもつといえよう。なおかつ多様な楽曲およびフレーズに対しても性能の低下がないといえる。

図 7 の散布図はやや非線形な分布をしているように見取れる。この原因として、主観評価値の分布と特徴量の分布との形状が異なることが考えられる。解決方法には back-end での評価値算出手法を何らかの非線形手法に置き換えることが考えられ、今後検討すべき課題のひとつである。

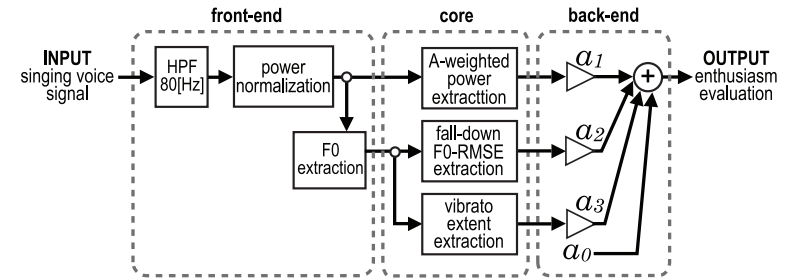


図 6 熱唱度評価システム SEES のブロック図
Fig. 6 Block-diagram of the SEES

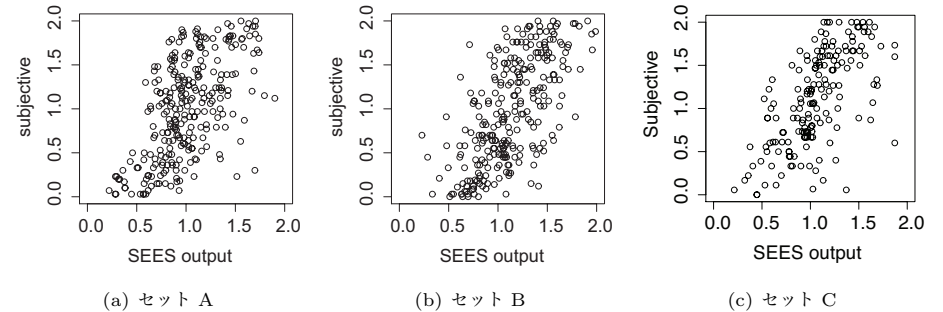


図 7 熱唱度評価システム SEES の出力と主観評価値
Fig. 7 Comparison of the SEES output and the subjective evaluations

6. iPhone アプリとしての実装

極めて手軽なエンタテインメントとして歌唱を楽しむことを目的に、SEES を図 8 のような iPhone アプリとして実装した。iPhone は底面にマイクを搭載しており、図 9 のように iPhone 本体をハンドマイクに見立てて歌唱することができる。操作ボタンとして歌唱の開始/終了ボタンと再生ボタンの 2つのみを搭載するシンプルなインターフェースとした。カラオケのように大掛かりな装置を必要とせず、また選曲等の煩わしい操作をすることなく、歌うことそのものを楽しむ新しい形態のアマチュア歌唱エンタテインメントとして提案する。



図 8 iPhone アプリとしての SEES 実装
Fig. 8 An SEES implementation for iPhone



図 9 iPhone をハンドマイクに見立てて歌う
Fig. 9 Singing with iPhone as like as a microphone

7. むすび

本稿に述べたことを以下にまとめる。

熱唱度評価システム SEES を開発するにあたり、まず主観評価実験の結果から、熱唱度の知覚に関して以下のことを示した。

- 人間は歌唱音声の聴取のみによって、おおよそ共通の基準で熱唱度を知覚できる
 - 本人熱唱度・フレーズ・歌唱者のいずれもが知覚熱唱度に対して有意な効果をもつ
- また熱唱度の自動評価に関して、以下のことを確認した。
- A 特性パワー・ずり下ろし・ビブラート深さの 3 つの特徴量が、フレーズによらず自動評価に有効
 - 上記特徴量の重み付き線形和による自動評価で、主観評価値との間に 0.6 以上の相関
 - 特徴量を非線形に組み合わせることで、さらに相関が向上する可能性

さらに SEES を用いた手軽なアマチュア歌唱エンタテインメントのアプリケーションとして、iPhone への実装について述べた。

本研究では短い歌唱フレーズのみを用いてきたので、今後は 1 曲全体などのより長い歌唱音声に対して、熱唱度の知覚に関する分析や自動評価手法の検討を進めたい。

参 考 文 献

- 1) Cano, P., Loscos, A., Bonada, J., de Boer, M. and Serra, X.: Voice Morphing System for Impersonating in Karaoke Applications, *Proc. ICMC* (2000).
- 2) Jansens, S., Bloothoof, G. and de Krom, G.: Perception and Acoustics of Emotions in Singing, *Proc. Eurospeech*, pp.2155–2158.
- 3) Justlin, P.N. and Laukka, P.: Communication of Emotions in Vocal Expression and Music Performance: Different Channels, Same Code?, *Psychol. Bull.*, Vol.129, No.5, pp.770–814 (2003).
- 4) Kotlyar, G.M. and Morozov, V.P.: Acoustical correlates of the emotional content of vocalized speech, *Soviet Physics Acoustics*, Vol.22, No.3, pp.208–211 (1976).
- 5) Lal, P.: A Comparison of Singing Evaluation Algorithms, *Proc. ICSLP* (2006).
- 6) Mayor, O., Bonada, J. and Loscos, A.: Performance Analysis and Scoring of the Singing Voice, *Proc. 35th AES Intl. Conf.*, pp.1–7 (2009).
- 7) Mitsui, T. and Hosokawa, S.: *Karaoke around the world*, Routledge (1998).
- 8) Nakano, T., Goto, M. and Hiraga, Y.: An Automatic Singing Skill Evaluation Method for Unknown Melodies Using Pitch Interval Accuracy and Vibrato Features, *Proc. ICSLP* (2006).
- 9) Scherer, K.R.: Expression of Emotion in Voice and Music, *J. Voice*, Vol.9, No.3, pp.235–248 (1995).
- 10) Siegwart, H. and Scherer, K.R.: Acoustic Concomitants of Emotional Expression in Operatic Singing: The Case of Lucia in *Ardi gli incensi*, *J. Voice*, Vol.9, No.3, pp.249–260 (1995).
- 11) Sjölander, K.: The Snack Sound Toolkit, The Department of Speech, Music and Hearing, KTH (online), available from <http://www.speech.kth.se/snack/> (accessed 2011.1.6).
- 12) Tsai, W.-H. and Lee, H.-C.: An Automated Singing Evaluation Method for Karaoke Systems, *Proc. ICASSP*, pp.2428–2431 (2011).
- 13) 大道竜之介, 伊藤 仁, 伊藤彰則, 牧野正三: 歌唱の「熱唱度」評価の検討, 情報処理学会研究報告 MUS-85 (2010).
- 14) 森瀬将雅, 中野皓太, 西浦敬信: 実時間歌唱力補正に基づく新たなカラオケエンタテインメントの創出, 情報処理学会研究報告 MUS-86 (2010).
- 15) 公益財団法人日本生産性本部: レジャー白書 (2011).