

リアルタイム日本語歌唱鍵盤楽器 “VOCALOID キーボード” の開発

加々見 翔太[†] 濱野 桂 三[†]
柏瀬 一輝^{††} 山本 和彦^{†††}

本稿では、リアルタイムに日本語の歌詞とその音高を同時に入力して、歌唱合成を利用した演奏をすることを可能とするリアルタイム日本語歌唱鍵盤楽器を提案する。本システムのボタン配列を用いた歌詞入力インターフェースでは、ローマ字入力方式とかな入力方式を組み合わせることによって非常に少ない数のボタンのみを利用して任意の日本語の文字を高速に入力することが可能である。この歌詞入力インターフェースは片手で操作可能のようにボタン配置が最適化され、一般的な12音階の鍵盤を音高入力装置として同時に組み合わせることで、歌唱合成に必要なパラメータの入力をリアルタイムに行うことができる。また、本システムは演奏技術の習得が非常に容易であり、検証でもその有効性が確認された。

Development of Realtime Japanese Vocal Keyboard.

SHOTA KAGAMI,[†] KEIZO HAMANO,[†] KAZUKI KASHIWASE^{††}
and KAZUHIKO YAMAMOTO ^{†††}

In this work, we present an instrument called "Realtime Japanese Vocal Keyboard" that enables us to perform music using singing synthesis in realtime by inputting japanese lyrics and the pitch simultaneously. In this system, the interface of button matrix for inputting lyrics, which combines "romaji" method with "kana" method, enables us to input arbitrary japanese characters using few combination of buttons very quickly. This lyric input interface is optimized for single hand use, and by combination with common twelve-tone musical keyboard, we can input the parameters need for singing synthesis in realtime. Additionally, this system don't requires hard training, very easy to learn, and the experiments show the effectivity.

1. はじめに

現在の音楽文化を取り巻く状況において、UGC(User Generated Contents), CGM(Consumer Generated Media) と呼ばれる分野は非常に重要な位置を占めるようになってきている。その中でも一般のユーザやアーティストが VOCALOID¹⁾²⁾ を始めとした歌唱合成ソフトウェアを用いて楽曲制作を行い、Web 上で発表をすることが盛んに行われており、歌唱合成ソフト

ウェアというものが我々にとって、より身近な存在として認知されるようになってきている。こうした歌唱合成ソフトウェアを使うことで、歌を歌うことが得意でないユーザでも自分の楽曲に歌唱を導入することができ、コンピュータで音楽制作を行う多くのユーザにとって魅力的なツールとなっている。

従来の歌唱合成ソフトウェアで音楽を制作するための手順には現状、歌詞と音高をそれぞれ入力する作曲プロセス、それをレンダリングして再生する再生プロセスの2段階のプロセスを必要とする。これは特に歌詞と音高を同時に素早く入力することが困難であるということに起因し、このため、リアルタイムには意図した歌唱演奏を行うことはできないという問題がある。こうしたリアルタイムに歌唱合成を利用した演奏が容易に可能となれば、今までオフラインでしか歌唱合成を利用した音楽活動を行うことができなかったユーザがオンライン上で直接音楽を生演奏できるようになったり、NETDUETTO³⁾ といったネットワークを利用したリアルタイムセッションにおいて歌を歌え

[†] ヤマハ株式会社 デジタル楽器事業部 商品開発部 ソフト開発グループ

Software Design Group Product Development Department, Digital Musical Instruments Division, Yamaha Corp.

^{††} ヤマハ株式会社 デザイン研究所

Design Laboratory, Yamaha Corp.

^{†††} ヤマハ株式会社 デジタル楽器事業部 技術開発部 システム開発グループ

Technology Development Group Technological Development Department, Digital Musical Instruments Division, Yamaha Corp.

ないユーザが歌唱合成演奏を導入したりできるようになり、音楽表現の方法に新たな自由度を与えることができ有用である。

そこで本稿では、リアルタイムに日本語の歌詞とその音高を同時に入力して歌唱合成を利用した演奏することを可能とするインターフェースを構築する。本システムでは、ボタンを使ったインターフェースにローマ字入力とかな入力の特徴を組み合わせた日本語入力システムと、広く普及している鍵盤を組み合わせることで、ユーザにとってより操作しやすいリアルタイム日本語歌唱のための鍵盤楽器の開発を目指した。本稿ではその途中経過を報告する。ここで、対象を日本語に限定したのは、他の多くの言語が全ての文字を平等に扱うのに対し、日本語では母音と子音が完全に分離しているために文字の扱いが平等ではないため、この性質を有効に利用した入力手法の構築が期待できると考えたためである。また、日本語はモーラ言語であることから、ユーザの入力した歌詞と音高の対応付けをする上で曖昧さが少ないという特徴もあり、リアルタイムの歌唱合成インターフェースと相性が良いと判断した。

2. 関連研究

歌唱合成を実現するためのデータ入力方法としては、歌詞と音高をあらかじめ入力しておくという方法以外にも過去にいくつかの研究が提案されている。Vocal-istener⁴⁾では、人間が実際に歌唱した歌データから自動的に抽出した特徴ベクトルを声色のベクトル空間にマッピングを行い、歌を歌えば自動的に VOCALOID の歌唱が出力される。しかし、歌唱した結果をリアルタイムに分析する事はできず、生演奏には適用できないため、本研究の目的のためには利用できない。また予測変換や機械学習などによる文字入力の補佐システムを利用するというアプローチがあり、マイクを用いて歌唱によって歌詞認識を行う研究⁵⁾や、手に装着するグローブ型デバイスを利用した指の形による日本語入力⁶⁾、カメラを使った指モーション認識⁷⁾などがあるが、これらは認識精度の点で解決しなければならない問題が多い。竹川らによる小型鍵盤を用いた文字入力インターフェース⁸⁾では、鍵盤のみを用いて日本語のリアルタイム入力を実現している。これは連続して複数の鍵盤から分散和音を入力して、その基音となる音からの他の音の音程とその順番を利用して鍵盤入力のみで語彙を入力する方法であり、鍵盤演奏の技術を有する12音階の鍵盤に慣れたユーザにとっては文字列を素早く入力することが可能である。しかし、リア

ルタイムに歌唱合成を利用した演奏を行うためには任意のリズムに合わせた入力が必要であり、こうした分散和音型の入力は、複数の鍵を順番に押すためにユーザにとってタイミングの制御が難しいという問題があり、本稿の目的には適わない。ここで、加藤ら⁹⁾¹⁰⁾による日本語入力をテンキーで行う試みは、ボタンのみを使う構成であり、実現が容易な他、母音キーと子音キーが分離し、その組み合わせで入力するという点で、コンピュータのキーボード上でのローマ字入力に慣れたユーザなら比較的理解されやすいインターフェイスであるという利点がある。また片手で入力できるので、もう一方の手で音高の指定操作を行えるなど、歌詞を入力しながら何らかの楽器演奏をするということと親和性が高いと考えられる。加えて所望の文字を入力するためにボタンを押す順番を意識しなくてもよいので、発音のタイミング制御も容易である。そこで本稿では、加藤ら⁹⁾の手法を発展させ、歌唱入力と音高指定を同時に行うことのできるインターフェースを開発する。

3. 提案手法

本稿では、日本語入力をボタン、音高を鍵盤で入力し、その場でリアルタイムに楽器に歌わせる・演奏する事を目的としたリアルタイム日本語歌唱鍵盤楽器”VOCALOID キーボード”(図1)を開発する。

リアルタイムに日本語を歌唱させる事ができる鍵盤楽器を実現するためには、

- 高速な日本語入力インターフェース
- 音高入力インターフェース
- 歌唱の抑揚や声質を制御するインターフェース
- それらを処理し、合成歌唱を出力するシステム

がそれぞれ必要である。本システムは、これらの要求を満たすため、加藤ら⁹⁾の研究を基にしたリアルタイム日本語入力手法と、広く普及し利用者も多い鍵盤インターフェースを組み合わせた。また、歌唱の抑揚や声質を制御するインターフェースとしてのツマミインターフェース、合成歌唱を出力する音源として VOCALOID-board(後述)を採用した。実際の入力方法としては、ユーザは任意の文字を表現するボタンの組み合わせを日本語入力インターフェース上で左手で押さえて入力した上で、右手の音高入力インターフェースで音高を指定し発音を行う。以下その詳細について説明する。



図 1 VOCALOID キーボード
Fig.1 VOCALOID Keyboard



図 2 日本語入力インターフェース (文字が書いてあるボタンのみ利用)

Fig.2 The japanese lyrics input interface

3.1 日本語入力インターフェース

本システムでは目的とする歌詞をできるだけ素早く入力できるように、加藤ら⁹⁾による日本語入力手法を改良して採用した。このインターフェースでは入力方法の複雑さを最小限に抑えるため、出来るだけ少ないボタン数で日本語の発音記号を満たすように子母音を16キーに絞り込んだ。母音ゾーンと子音ゾーンは分離して配置している(図2)。これは、コンピュータのキーボードでのローマ字入力方式に慣れたユーザには子音と母音の組み合わせという入力方法は理解しやすく、また高速に入力可能であると考えたためである。ホームポジションとしては左手を置いた際に親指近くに母音、その他の指に子音が近くなるような配置にした。親指で5母音、それ以外の指で各指2~3子音+濁音・半濁音を担当する。基本的にはローマ字入力方式を踏襲しており、子音+母音を同時に押す事で目的とする文字を入力する。通常のローマ字入力方式

と異なる点としては、濁音と半濁音には別途濁点ボタンと半濁音ボタンを用意し、濁音と半濁音入力時には、これを同時に押すこととした(例 ベ = h + " + e)。そのためローマ字入力における B, D, V, G, Z, P といったボタンの削減を可能にしている。これはかな入力の濁音・半濁音を入力において、濁点・半濁点を入力する方法を踏襲しており、ボタンの数を削減することで入力スピードを向上させる効果がある。また、本入力システムの特徴としては、母音が"あ"である発音は、"A"ボタンの母音入力をしなくても、子音のみ押した状態で発音が可能である。例外としては、日本語で頻出する"ん"の入力は"n"のみで行えるようになっている。入力が複雑になりやすい促音に関しては、複数の入力方法やローマ字入力使われにくい組み合わせのショートカット入力を用意している(例: みゅ 通常 m + y + u / ショートカット m + h, びゃ 通常 h + ° + y + a / ショートカット h + °)。このような入力方法は本方式に慣れたユーザがさらに素早く入力する際に利用できる。押下ボタンと発音文字列の対応に関しては(図7)に詳細を載せた。このようにローマ字入力とかな入力を組み合わせ、最小限の押下にて目的とする発音を行うことを可能にすることで、入力の速度・リアルタイム性の向上を狙った日本語入力インターフェースを考案した。

3.2 音高入力インターフェース

本稿では、音高の入力装置として一般的に普及している12音階の鍵盤を利用する。これはユーザにとって演奏技術の習得を容易する目的の他、提案する歌詞入力インターフェース単体の有効性の検証を行うためには、音高入力装置において既に多くの被験者が演奏技術を持ち合わせているインターフェースを採用することが適していると判断したためである。実際の発音タイミングは鍵盤が押されたタイミングとなり、このときに歌詞入力インターフェースにおいて左手で押さえ



図 3 入力文字表示用 Led マトリックス
Fig. 3 Led matrix display for Input words

られていたボタンに対応した歌詞を発音する。これは日本語がモーラ言語であるからこそ可能なアプローチである。

3.3 文字表示システム

文字表示システム (図 3) は、現在発音している文字列と既に過去に発音した文字列を表示する。従来の楽器では演奏者自身へのフィードバックが音のみで行われる事が多いが、歌唱合成を使ってリアルタイムに演奏する上で特有の問題である発音された歌詞のフィードバックを確実にを行うためにこうした文字列の表示が必要である。

これは、後述する音源である VOCALOID-board の歌詞の発音は明瞭でないこともしばしばあり、ユーザへのフェードバックとして音だけでは分かりにくい場合があるため、それを補助する役割を持っている。また、ユーザが歌詞入力インターフェースにおいてブラインド入力を練習する際に、自分の入力した文字の正誤が判断できなくてはならない。そのため自分の入力を確認するために必要となる表示装置でもある。この表示システムでは、現在入力している文字が点滅することで、発音済みの文字と入力中の文字の違いを表現している。また、演奏後の文字列を表示する事によって、演奏者が正確に演奏できているかを認識する事ができるようになる。

3.4 歌唱合成エンジンとパラメーター操作

本稿では、歌唱合成のエンジンとして VOCALOID-board (図 5) を利用する。これは MIDI のシステムエクスクルーシブを入力することにより、リアルタイムに歌唱合成を行うことのできる組み込みボードである。歌唱合成をするためのパラメータは歌詞と音高以外にも数種類存在するが、その中でも以下にある歌声パラメーター、

- VIBRATO DEPTH : ビブラートの深さ
- VIBRATO RATE : ビブラートの速さ
- PORTAMENTO TIMING : ポルタメント



図 5 VOCALOID-board
Fig. 5 VOCALOID Board

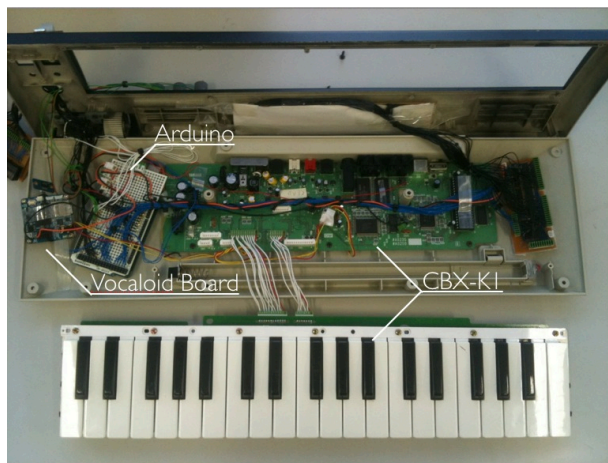


図 6 システム
Fig. 6 The System

- BREATHNESS : 息づかい
- BRIGHTNESS : 明るさ
- CLEARNESS : 明瞭さ
- GENDER : 性別

をツマミによって変更できるようにし、自分なりの声質を作り込む事が可能とした。

4. 実 装

提案手法の検証のためにプロトタイプ (図 4) の開発を行った。システムの I/O 処理に ArduinoMega2560¹¹⁾、日本語入力インターフェイス部に tkrworks パッドキット¹²⁾、鍵盤部分に CBX-K1¹³⁾、文字列表示部分に 8 × 8 マトリックス LED × 7 個、実際の歌唱の音声合成を行う装置として VOCALOID-



図 4 プロトタイプ
Fig. 4 The prototype

board(図 5) を利用した。

動作の流れとしては、日本語入力インターフェースから 1 発音分の文字列を入力し、鍵盤で音高を指定、それらの信号を Arduino 上で VOCALOID-board へ入力できるように MIDI のシステムエクスクルーシブに変換し送信、VOCALOID-board 上で音声を合成し出力する。LED への発音文字の表示やツマミからの入力も Arduino からの制御で行なっている。

5. 検 証

提案手法を検証するため、実際に幾人かの被験者に演奏を体験してもらいその様子を観察した。いずれの場合も特に奏法などの指示は行わず、自由に演奏してもらった。歌詞入力しながら鍵盤による音高入力を同時に行うため、鍵盤楽器の演奏技術を全く有さない被験者であると、鍵盤と歌詞入力インターフェースのそれぞれの有用性を分離して検証できない。そのため今回は鍵盤楽器の演奏経験の有する被験者を対象とした。

まず、鍵盤楽器の演奏歴が 10 年の被験者に本システムを使って練習・実演奏を行ってもらった結果、3 時間程度の練習で日本語入力が無理なくできるようになり、最終的には、簡単な童謡をスムーズに演奏できるようになった。日本語入力インターフェースに関しては、ルールなどを教える事も無く演奏しながら習得できていた。また、鍵盤楽器の演奏歴が 3 年の被験者にも体験してもらった結果、特に指示をする事も無く演奏を始めることができ、最初に触れてから 2 時間程で、最終的に一曲オリジナル曲を作曲し演奏する事が可能になった。歌詞入力のリアルタイム性についても、あまりテンポの速い曲を演奏するのでもなければ十分に対応できており、さらなる演奏者の習熟によってこの点は向上する可能性がある。このことから、ある程度の鍵盤演奏経験があれば本システムは特に習熟が困難でなく、自由に歌詞を入力しながら歌唱合成を使った演奏を行うことができ、提案手法に有効性があること

が分かった。

今後は鍵盤ユーザーと被鍵盤ユーザーの両方を対象としてテストするなど、より多くの実験、検証を通して、提案手法のさらなる有効性を確認していく。

6. まとめ・今後の課題

本稿では、リアルタイムに歌詞と音高を同時に入力して歌唱合成を使った演奏を行うことを可能にするインターフェースとしてリアルタイム日本語歌唱鍵盤楽器“VOCALOID キーボード”のプロトタイプ開発を行った。また、試作を数人の被験者に対して実際に演奏してもらいその有効性を確認した。今後の課題としては、より高速な歌詞入力を可能にするための改良を行っていく。また、歌唱のパラメータは今回の手法ではつまみで行うため、歌詞と音高入力で両手が塞がれている状態では操作することができないという問題がある。こうしたパラメータも含め、強弱、抑揚等の歌唱の微妙なニュアンスを歌詞と音高を入力すると同時にリアルタイムに演奏するインターフェースを今後構築していく。

参 考 文 献

- 1) ヤマハ株式会社：“VOCALOID”
<http://www.VOCALOID.com/>.
- 2) 剣持 秀紀：“歌声合成とその応用”，情報処理 (Vol.50 No.8 Aug.2009).
- 3) ヤマハ株式会社：“NETDUETTO”，
<http://www.y2lab.com/project/netduetto/>
- 4) 中野 倫靖，後藤 真考：“VocaListener：ユーザ歌唱を真似る歌声合成パラメータを自動推定するシステムの提案”，情報処理学会 音楽情報科学研究会 研究報告 (Vol.2008,No.50,pp.49-56,May 2008) .
- 5) 伊藤 直樹，西本 一志：“歌詞歌唱による入力可能な Voice-to-MIDI 手法の提案”，情報処理学会シンポジウム論文集,(2007.4.71-72.)
- 6) 福島 大志，宮崎 文夫，西川 敦：“指文字入力イ

わ	ら	や	ま	は	な	た	さ	か	あ
w / wa	r / ra	y / ya	m / ma	h / ha	na	t / ta	s / sa	k / ka	a
うい	り	ゐ	み	ひ	に	ち	し	き	い
wi	ri	yi	mi	hi	ni	ti	si	ki	i
う	る	ゆ	む	ふ	ぬ	つ	す	く	う
wu	ru	yu	mu	hu	nu	tu	su	ku	u
うえ	れ	いえ	め	へ	ね	て	せ	け	え
we	re	ye	me	he	ne	te	se	ke	e
を	ろ	よ	も	ほ	の	と	そ	こ	お
wo	ro	yo	mo	ho	no	to	so	ko	o

ん
n

ば	ば
h [˙]	h [˙]
び	び
h [˙] i	h [˙] i
ぶ	ぶ
h [˙] u	h [˙] u
べ	べ
h [˙] e	h [˙] e
ぼ	ぼ
h [˙] o	h [˙] o

だ	ざ	が
t [˙] / t [˙] a	s [˙] / s [˙] a	k [˙] / k [˙] a
でい	じ	ぎ
t [˙] i	s [˙] i	k [˙] i
づ	ず	く
t [˙] u	s [˙] u	k [˙] u
で	ぜ	げ
t [˙] e	s [˙] e	k [˙] e
ど	ぞ	こ
t [˙] o	s [˙] o	k [˙] o

りゃ
ry / ria
りゅ
ryu / rh / nw / riu
りえ
rie / rm
りょ
ryo / rho / rw / rio

みゃ	ひゃ	にゃ	ちゃ	しゃ	きゃ
my / mia	hy / hia	ny / nia	ty / tia	sy / sia	ky / kya / kha / kma / kwa / kia
みゅ	ひゅ	にゅ	ちゅ	しゅ	きゅ
myu / mh / mwu / miu	hyu / hwu / hiu	nyu / nh / nwu / niu	tyu / th / twu / tiu	syu / sh / swu / siu	kyu / kh / khu / kmu / kwu / kiu
みえ	ひえ	にえ	ちえ	しえ	きえ
mie	hm / hie	nm / nie	tm	sm / sie	kye / khe / km / kme / kwe / kie
みょ	ひょ	にょ	ちょ	しょ	きょ
myo / mho / mw / mio	hyo / hw / hio	nyo / nho / nw / nio	tyo / tho / tw / tio	syo / sho / sw / sio	kyo / kho / kmo / kw / kwo / kio

てい
tei

びゃ	びゃ	ふぁ	ぢゃ	じゃ	ぎゃ	づぁ
h [˙] y / [˙] / ia	h [˙] y / [˙] / ia	hua	t [˙] y	s [˙] y / s [˙] ia	k [˙] y / k [˙] ha / k [˙] ma / k [˙] / wa / k [˙] ia	u [˙] a
びゅ	びゅ	ふい	ぢゅ	じゅ	ぎゅ	づい
h [˙] yu / [˙] u / iu	h [˙] yu / [˙] u / iu / h [˙] iu	hui	t [˙] yu / t [˙] h / t [˙] wu	s [˙] yu / s [˙] h / s [˙] wu / s [˙] iu	k [˙] yu / k [˙] h / k [˙] wu / k [˙] iu	u [˙] i
びえ	びえ	ふえ	ぢえ	じえ	ぎえ	づえ
[˙] m / [˙] ie	[˙] m / [˙] ie / h [˙] / ie	hue	t [˙] m	s [˙] m / s [˙] ie	k [˙] m / k [˙] ie	u [˙] e
びょ	びょ	ふぉ	ぢょ	じょ	ぎょ	づょ
h [˙] w / [˙] o / io	h [˙] w / [˙] o / io / h [˙] io	huo	t [˙] yo / t [˙] / ho / t [˙] w	s [˙] yo / s [˙] h / ho / s [˙] w / s [˙] io	k [˙] yo / k [˙] h / ho / k [˙] w / k [˙] io	u [˙] o

図7 ボタン-かな文字 対応表

Fig. 7 The relation between buttons and kana

- ンターフェイス「Fingual」の開発”, 情報処理学会インタラクシオン 2011.
- 7) 西田 好宏, 鈴木 雅実: “VecType: 動きベクトルを利用した携帯電話文字入力システムの提案”, 情報処理学会インタラクシオン 2005.
 - 8) 竹川 佳成, 寺田 努, 塚本 昌彦, 西尾 章治郎: “歌唱機能をもつ黒鍵追加型小型鍵盤楽器モバイルクラヴィア IV の設計と実装”, 情報処理学会研究報告 (音楽情報科学研究会 2004-MUS-57, Vol.2004, No.111, Oct. 2004)
 - 9) 加藤 善也, 北上 義一: “複合入力処理によるパーム型キーボードの実験”, ヒューマンインタフェー

- ス 60-3 (1995.5.17) .
- 10) 加藤 善也, 北上 義一, 酒井 靖夫, 小野 充一: “テンキー型片手打鍵文字入力方式の実験と評価”, ヒューマンインタフェース 82-1 (1999.1.29) .
 - 11) Arduino: “Arduino Mega 2560”, <http://arduino.cc/>.
 - 12) (有) 山本製作所: “PICnome”, <http://atelier.tkrworks.net/doc/about-picnome>.
 - 13) ヤマハ株式会社: “CBX-K1”, <http://jp.yamaha.com/products/music-production/midi-controllers/cbx-k1/?mode=model>