

楽器音を用いた音声合成実現のための音素 MIDI コード設計機能の MIDI 符号化ツール「オート符」への実装

茂出木 敏雄[†]

[†] 大日本印刷株式会社 情報コミュニケーション研究開発センター
〒141-8001 東京都品川区西五反田 3-5-20

あらまし 筆者らは、音声信号に対して高精度な和音解析を行い、フォルマント成分を和音近似し、MIDI 符号に変換することにより、一般的な MIDI 音源を用いて音声再生する技術を確認した。その後、和音演奏可能な玩具楽器や、アコースティック楽器により、音声再現の要望が出てきたが、既開発の MIDI 符号化ツールで生成される MIDI データは標準 32 和音で構成され、木目細かい演奏指示情報が付与されているため、判読性のある五線譜に変換することは困難であった。本稿では、カナテキスト入力により MIDI 音源や楽器演奏により音声再生できる MIDI データを合成するシステムを開発するため、日本語の各音素を二連の単純な和音で表現した MIDI コードのデータベースを設計する機能を既開発の MIDI 符号化ツール「オート符」に実装したので、その概要を紹介する。

Implement of Phoneme MIDI Code Designing Functions on MIDI Encoding Tool “Auto-F” for Realizing Voice Synthesizing Based on Musical Sounds

Toshio MODEGI[†]

[†] Media Technology Research Center, Dai Nippon Printing Co., Ltd.
3-5-20, Nishigotanda, Shinagawa-ku, Tokyo 141-8001, Japan
E-mail: [†] Modegi-T@mail.dnp.co.jp

Abstract We have developed a vocal synthesizing technique using a general MIDI sound modules by converting given vocal signals into MIDI harmonic codes, which represent analyzed formant components. Then vocal signal simulations using toy or acoustic musical instruments have been required. However, it is difficult to convert encoded MIDI codes to readable music scores, because they are normally composed of 32 harmonic tones with complex articulation instruction codes. In this paper, we propose constructing vocal MIDI code synthesizing system, which can convert given Japanese kana-texts into MIDI data stream to be played back with MIDI sound modules or acoustic musical instruments. We present our implemented phoneme MIDI code designing functions, which can represent each phoneme with a pair of simple harmonic codes, on our previously developed MIDI encoding tool.

1. まえがき

筆者らは与えられた音響信号に対して一般化調和解析を用いて平均律音階のスケールで高精度な周波数解析を行い、MIDI データ形式に自動変換する技術の開発を進めてきた[2]-[4]。本技術は「オート符@SA」という名称で汎用的な音響解析ツールとしてまとめ、2001 年より財団法人デジタルコンテンツ協会のホームページより無償配布を進めており、主として採譜業務の支援等に活用いただいている[5]。本解析ツールは、特に和音解析精度が高く、音声信号に適用すると解析されたフォル

マント成分が MIDI 形式に和音近似され、一般的な MIDI 音源を用いてボーカルが再現できるという特徴をもつ[1]。

これに対して、その後、和音演奏可能な玩具楽器や、複数のアコースティック楽器等をアンサンブル演奏することにより、ボーカルが再現できないかという要望が多く寄せられるようになり、技術的な検討を進めてきた。まず、既開発の MIDI 符号化ツールで生成される MIDI データは標準 32 和音で構成され、木目細かいデュレーションやベ

ロシティなどの演奏指示情報が付与されているため、そのままでは判読性のある五線譜に変換することは困難であった。検討の結果、日本語音素については、デュレーションやペロシティを均一にした二連の単純な和音で表現する方法を検討し、カナテキスト入力により MIDI 音源や楽器演奏により音声を再現できる MIDI データを合成するシステムの試作に着手した。

本稿では、日本語 50 音の各音素を二連の単純な和音で表現した MIDI コードのデータベースを構築するため、各音素を録音した波形音響信号より自動変換された高精細な MIDI データに対して、二連の単純な和音に抽象化する機能を既開発の MIDI 符号化ツール「オート符」に実装したので、その概要を紹介する。

2. 既提案の音響信号の MIDI 符号化ツールのアルゴリズム概要

図 1 は筆者らが開発した MIDI 符号化処理の主要構成を示す[4]。はじめに、与えられたソース音響信号より周波数解析対象のフレームを抽出するが、後続フレームへのシフト幅はソース音響信号の周波数変動を大まかに検出しながら適応的に設定するようにしている。続いて、平均律音階の半音（ノートナンバー）単位に非線形な周波数次元で周波数解析を行うが、周波数が高くなるにつれ、半音間隔が粗くなるため、周波数ごとに半音間を微分音（副周波数）に分割して解析を行うようにしている。最後に、時間的に隣接する同一主周波数の解析成分（音素）を連結し音符としてまとめ、MIDI イベント形式で符号化する。図 1 例では、単一の MIDI イベントで記載しているが、実際には

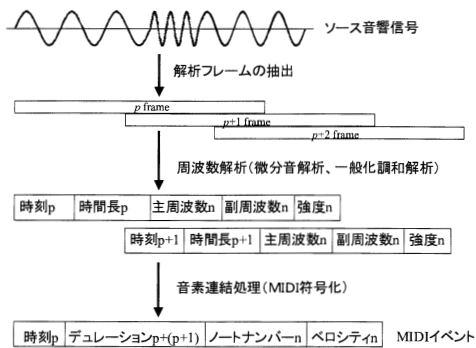


図 1 提案する MIDI 符号化ツールの主要構成

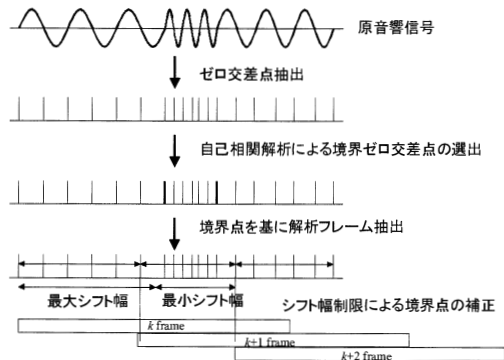


図 2 解析フレーム抽出処理の説明

デュレーション情報がなく時刻が異なる 2 つのノートオンとノートオフ・イベントで符号化される。

2.1. 解析フレームの抽出

音響解析では解析フレームを時間軸方向に移動させながら、信号全体の解析を行うが、本節ではこの際のフレーム長とシフト幅の設定方法について述べる。

周波数分解能はフレーム長により変化し、経験上ソース音響信号のサンプリング周波数が 44.1[kHz] の場合、低域部まで忠実に解析するためには 4096 サンプル以上必要である。フレーム長は固定値でよいと、例えば $T=4096$ とする。

一方フレームシフト幅は、小さくするほど時間分解能が向上するが計算時間も増大する。そして、解析対象信号が単調である箇所に対して、必要以上にフレームシフトを細かくすると、後述する音素連結処理で支障をきたす。そこで、効率的な計算および高精度な音素連結処理のためにも、フレームシフト幅は解析対象フレームごとに変化させ、最適な値を設定する方法が望ましい。

図 2 は筆者らが提案する最適な解析フレーム位置を設定する手法を示す。始めに、与えられた音響信号の全領域に対してゼロ交差点検出を行い、ゼロ交差点間隔の粗密または自己相関解析により周波数変化が顕著なゼロ交差点を選別する。基本的には、この変化が急峻なゼロ交差点に解析フレームを設定するが、例えば、最小シフト幅を $T/8$ に、最大シフト幅を $T/2$ に設定し、シフト幅が一定以上に細かくなり過ぎたり、延び過ぎたりしないように調整を行う。

2.2. 周波数解析

前節で述べた方法により、サンプリング周波数 f_s の原音響信号より p 番目に抽出された解析フレームのサンプル配列を $x(i)$ ($0 \leq i \leq T-1$) とする。周波数解析は、 n ($0 \leq n \leq 127$) を MIDI のノートナンバーとして 128 種の解析周波数 $f(n) = 440 \cdot 2^{(n-69)/12}$ の調和関数を基本にした一般化調和解析により行う。しかし、周波数が高くなるにつれ、ノートナンバー間の周波数間隔が広がるため、特に $n > 60$ では解析精度が低下してしまう。そこで、ノートナンバー間を以下のように M 個の微分音に分割した $128M$ 種の調和関数を用いて解析を行う。

$$f(n, m) = 440 \cdot 2^{(n-69+m/M)/12}. \quad (1)$$

はじめに、ノートナンバー分の強度配列 $E(n)$ ($0 \leq n \leq 127$) と副周波数配列 $S(n)$ を定義し、初期値を全て 0 とする。

(a) $0 \leq n \leq 127$ のおよび $0 \leq m \leq M-1$ に対して以下式で相関計算を行い、 $E(n, m)$ を最大にする (n_{max}, m_{max}) を求める。式(2)において $T(n)$ は解析フレーム長で、フレーム長 T を超えない範囲で調和関数の周期の最大の整数倍になるよう設定し、 k を適当な整数値として、 $T(n) = k/f(n, m)$ で与える。

$$\begin{aligned} A(n, m) &= \sqrt{T(n)} \cdot \sum_{i=0, T(n)} x(i) \sin(2\pi f(n, m) i / f_s) \cdot \\ B(n, m) &= \sqrt{T(n)} \cdot \sum_{i=0, T(n)} x(i) \cos(2\pi f(n, m) i / f_s) \cdot \\ E(n, m) &= A(n, m)^2 + B(n, m)^2. \end{aligned} \quad (2)$$

(b) 上記決定した $A(n_{max}, m_{max})$ および $B(n_{max}, m_{max})$ を用いて、以下式でサンプル配列 $x(i)$ の全ての要素 ($0 \leq i \leq T-1$) を更新する。

$$\begin{aligned} x(i) &= x(i) - A(n_{max}, m_{max}) \cdot \sin(2\pi f(n_{max}, m_{max}) i / f_s) \cdot \\ &\quad - B(n_{max}, m_{max}) \cdot \cos(2\pi f(n_{max}, m_{max}) i / f_s). \end{aligned} \quad (3)$$

(c) $E(n_{max}) \leftarrow E(n_{max}) + E(n_{max}, m_{max})$ 、 $S(n_{max}) = m_{max}$ として、再度(a)の処理に戻り、 $0 \leq n \leq 127$ の全ての $E(n)$ および $S(n)$ の値が決定するまで(a)から(c)までの処理を繰り返す。

処理負荷を軽減するため、 M の値についてはノートナンバーに基づいて可変に設定し、例えば解析する周波数間隔が 100 [Hz] 程度になるようにする。そして、ノートナンバー 60 以下は分割せず $M=1$ にする。また、精度は若干落ちるが、初回(a)の処理で $S(n)$ を決定し、2回目以降の(a)の処理は $m=S(n)$ に固定させ、微分音解析を省略する方法もとれる。

また、上記(a)の段階で、既に同一ノートナンバ

一に対して副周波数が異なる信号成分が複数回に渡って解析される可能性があるが、 $E(n)$ と $S(n)$ に既に値がセットされている場合は $E(n, m)$ の最大値の選定候補から除外する方法もとれる。

2.3. 音素連結処理

p 番目と $p+1$ 番目の解析フレームにより周波数解析されたノートナンバー n の音素を [時刻(p), 時間長(p), 主周波数 n , 副周波数 $S(n)$, 強度 $E(n)$] と [時刻($p+1$), 時間長($p+1$), 主周波数 n , 副周波数 $S(n)$, 強度 $E(n)$] とする。時間長(p) は解析フレームの第1サンプルの原音響信号上の絶対サンプルアドレスをサンプリング周波数で除算することで得られる。時間長(p) は時刻($p+1$) - 時刻(p) で、時間長($p+1$) は時刻($p+2$) - 時刻($p+1$) で与えられる。

時間的に隣接するこれら2つ音素に対して、副周波数を考慮した周波数の差が $1/2$ 半音未満で、双方の強度が所定のしきい値以上でかつ双方の強度の差が所定以下で両者の連続性が認められる場合、後続音素を前方音素に連結統合する。ただし、連結後の主周波数・副周波数・強度は大きい方を与え、時間長は双方の和、即ち時刻($p+2$) - 時刻(p) で与える。

同一ノートナンバーにおける時系列の連結処理は、不連続が認められるまで繰り返し行い、最終的に統合された [時刻(p), 時間長(p), 主周波数 n , 副周波数 $S(n)$, 強度 $E(n)$] に対して、2つの MIDI ノートイベントに変換する。時刻(p) で、ノートナンバー n のノートオン・イベントを発行し、ベロシティ値は $E(n)$ の最大値を E_{max} として、 $128 \cdot \{E(n)/E_{max}\}^{1/4}$ で与える。時刻については、Standard MIDI File では、直前イベントとの相対時刻 (デルタタイム) で与える必要があり、例えば、 $1/1536$ [sec] の単位に変換して与える。そして、時刻(p) + 時間長(p) $\cdot C$ で、ノートナンバー n のノートオフ・イベントを発行する。 C は $0 \leq C \leq 1$ の実数で、使用する MIDI 音源の音色に依存するが、MIDI 音源の余韻を考慮して早めにノートオフ指示をするためである。 $C=1$ にしても MIDI 音源の処理上は問題ないが、後続音と部分的に重なる場合がある。

MIDI 符号に変換する段階で、MIDI 音源で処理可能な同時発音数やビットレートについても考慮する必要がある。時間軸方向に発音期間中のノートイベントの個数が、例えば 32 和音を超えている場合は、ベロシティ値とデュレーション値 (時間長) の積で優先度を評価し、指定和音数以下にな

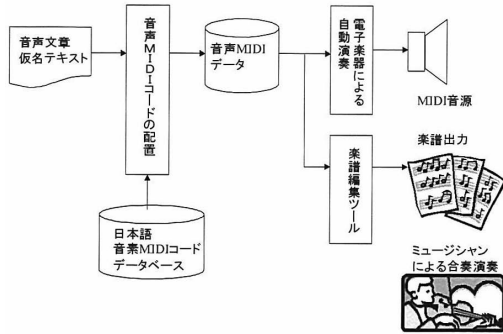


図3 楽器音を用いた音声合成システムの構想

るよう優先度の低いノートイベントを削除する処理を加える。また、デュレーションが短いノートイベントが多いと、ビットレートが増加しMIDI音源側の処理が間に合わなくなるため、同様に優先度の低いノートイベントを削除する。

3. 提案する楽器音を用いた音声合成システム構想と日本語音素のMIDIモデル

前節で述べた方法により、音声を録音した音響信号を与え、時間分解能 1/1536[sec], 和音数 32 の標準設定で MIDI 符号化を実行すれば、MIDI音源で音声を再生可能なMIDIデータを得ることができる。例えば、GM規格のMIDI音源でプログラム番号(54, Voice-Ooh)を設定して再生すれば、より音声らしい再生が可能である。

これに対して、既存の音声合成ツールと同様にテキスト入力で音声再生用のMIDIデータを出力できないか、和音演奏可能な玩具楽器や、複数のアコースティック楽器等をアンサンブル演奏することにより、ボーカルが再現ないかという要望が寄せられるようになってきた。即ち図3に示すような構想の実現性を検討した。前節で述べたMIDI符号化ツールでは、和音数が多く、木目細かいデュレーションやベロシティなどの演奏指示情報が付与されているため、そのままではヒトが演奏可能になるような判読性のある五線譜に変換することは困難であった。

そこで、日本語50音の音素について、デュレーションやベロシティを均一にし、できる限り単純な和音で音声として聴取可能な表現方法を模索した。音声は母音に主として現れる有声音と、子音に主として現れる無声音とに大別さ

れるが、各々スペクトル分解すると、前者は不連続なフォルマント・パターンが現れ、後者は連像的なノイズ・パターンになる。前者は4以上の音を重ねた和音で比較的容易に表現できるが、後者は非常に多くの音を重ねる必要がある。幸い本稿のように、MIDI音源の楽器音を重ねて再現する場合は、各構成音が倍音成分をもつため、比較的少

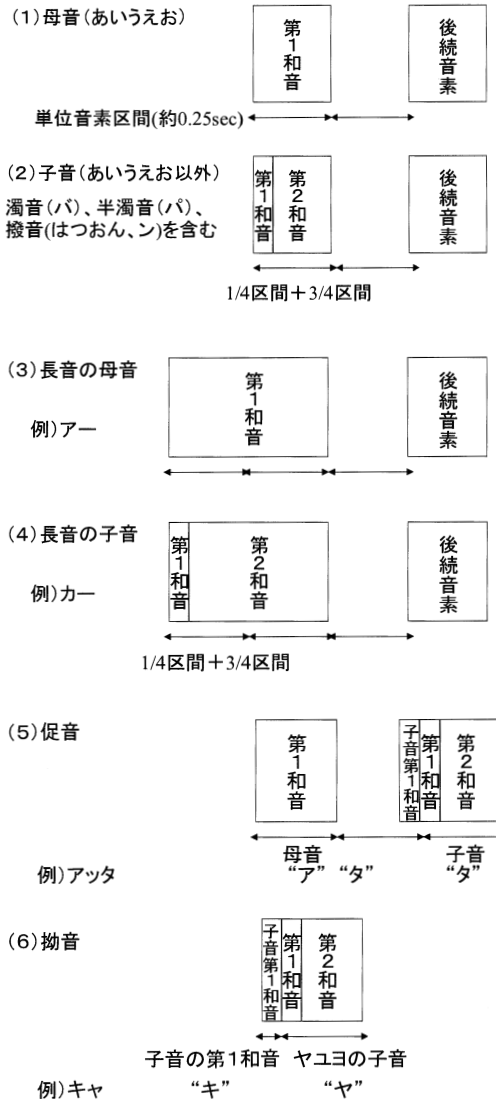


図4 提案する日本語音素の楽器音を和音近似モデル

ない音を重ねるだけで、ノイズ・パターンを近似することができる。また、日本語子音の場合は全て母音と連結するため、ノイズ・パターンからフォルマント・パターンに遷移する信号成分をもつが、これについては、視聴者の聴覚上の補間特性を活用することで、符号化の対象から外す。

図4に筆者らが提案する各種日本語音素を単純な和音で表現する方法について記す。基本構成は、同図(1)(2)に示されている通りで、あいうえおの母音のみ単一和音で表現し、それ以外の濁音、半濁音、撥音を含む子音類は全て二連の和音で表現する。和音数は原則4以上で最大8和音以下にする。ペロシティとデュレーションは均一にし、通常の母音音素の標準区間長を0.25[sec]にして、子音は第1和音を1/3区間長、第2和音を2/3区間長とする。

その他の長音・促音・拗音の扱いについては同図(3)~(6)に記す。促音・拗音のみ三連の和音構成になる。音声合成は、図3に示されるように、与えられた仮名テキストを基に前述した日本語音素MIDIコードのデータベースを参照し、所定の区間でMIDIデータを時間軸方向に結合した合成MIDIデータを作成すれば良い。その際、区間を変えれば話速を変更でき、MIDIデータを移調させれば、声の高さを変更でき、音素ごとに音楽のメロディーに合わせて移調させれば歌声にすることもできる。合成MIDIデータをMIDI音源で自動演奏させるか、既存の楽譜編集ツールで五線譜に変換すれば、ヒトによる楽器演奏でも音声合成を実現できる。

4. 提案する日本語音素 MIDI コード設計ツールの機能

前節で述べた日本語音素の和音モデルに基づきMIDIコードのデータベースを設計する手順を図5に記す。まず、母音および濁音、半濁音、撥音を含む子音を含む全71音の男声、女声の録音音声素材を収集する。

続いて、既開発のオーディオのMIDI符号化ツール「オート符」を用い、高精細なMIDIコードに自動変換する。変換結果例として、同図(A)に「カキクケコ」の音素サンプルに対して適用した結果を示す。横軸は時間で縦軸は周波数で、各四角形が音符(ノートオンとノートオフ・イベントの対)を示し、横幅はデュレーション、縦幅はペロシティの表示も兼用している。

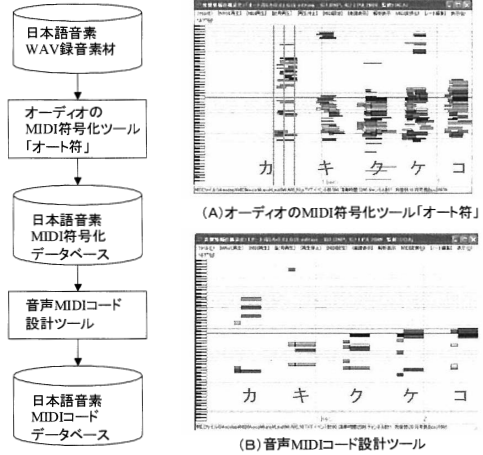


図5 日本語音素 MIDI コードのデータベース作成方法

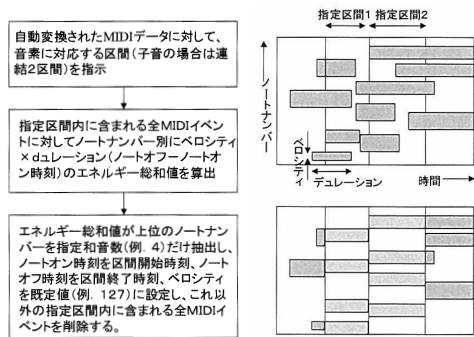


図6 音声 MIDI コード設計ツールの説明

前述の和音モデルに基づいて、音素ごとに単一または二連の8重音以下の和音に変換し、各音符のデュレーションおよびペロシティを均一化する。同図(A)のサンプルに対して、このような音符の整形化を行った結果が同図(B)である。本例では全ての和音に4和音に均一化し、ペロシティは全て127にする。母音のデュレーションと後続音素との間隔を各々標準区間長として0.25[sec]に設定し、子音は第1和音を1/3区間長、第2和音を2/3区間長とした。

図6に図5の音声MIDIコード設計ツールの具体的な機能の詳細を示す。MIDIデータに自動変換された単一音素に対応する区間を画面上で指示す

文 献

- [1] Toshio Modegi: "Multi-track MIDI Encoding Algorithm Based on GHA for Synthesizing Vocal Sounds," Journal of Acoustic Society of Japan, Vol.20, No.4, pp.319-324, April, 1999.
- [2] Toshio Modegi: "Very low bit-rate audio coding technique using MIDI representation," Proceedings of the ACM 11th international workshop on Network and operating systems support for digital audio and video, pp.167-176, June, 2001.
- [3] 茂出木敏雄, 「大サウンド特集: 音響情報のMIDI符号化ツール「オート符®」の開発」, 芸術科学会誌 DiVA (夏目書房), No.2, pp.42-48 (2002).
- [4] 茂出木敏雄: "音響信号の平均律音階に基づく汎用解析ツール「オート符」の開発," 電気学会・電子情報システム部門誌, Vol.123-C, No.10, pp.1768-1775, October, 2003.
- [5] 財団法人デジタルコンテンツ協会 d-CON Support, <http://www.dcaj.org/d-con/frame09.html> (「オート符®SA」の無償配布先)

る。母音の場合は、最初のイベントのノートオン時刻と最後のイベントのノートオフ時刻を設定すれば良いため自動化が可能であるが、子音の場合は2区間を指定する必要があり、境界部の時間位置は録音音素ごとに変化するため、対話形式に指示する方法をとった。

指定区間内で発音されているノートイベントのベロシティ値とでデュレーション値の積をエネルギー値として評価し、エネルギーが大きい順にノートイベントを指定和音数だけ選択し、残りは全て削除する。選択されたノートイベントに対しては、ベロシティを127に、デュレーションを指定区間になるように自動補正する。

5. 日本語音素 MIDI コード・データベースの制作例

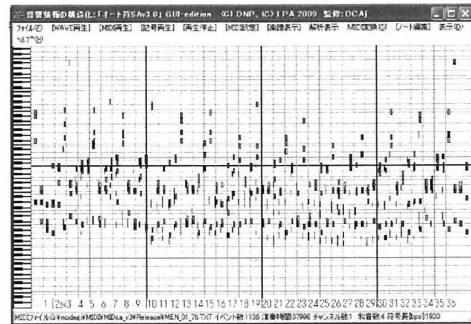
前節で述べた方法を用いて、母音および濁音、半濁音、撥音を含む子音を含む全71音の男声、女声の4和音構成による音素MIDIコードのデータベースを作成した結果を図7~9に示す。図8・9の各音名表記はMIDIノートナンバー60をC3とし、母音以外は2セットの4和音で構成され、第1和音と第2和音はつなげて演奏し、強さ・長さは均一であるが、第1和音は短めに、第2和音は長めに演奏することを想定している。

男声・女声の「アイウエオ」「カキクケコ」に対して市販の楽譜編集ツールを用いて五線譜に自動変換した結果、図10に示されるように判読性のある五線譜に変換できることを確認した。

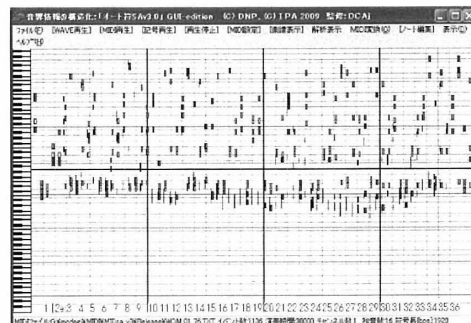
男声・女声の音素MIDIコードのデータベースを用いて「こんにちは」という仮名テキストを変換した結果を図11に示す。個々の音素は不明瞭ではあるが、GM規格のMIDI音源でプログラム番号(1, Grand-Piano)の設定でも音声らしきメッセージは聴取でき、音色をプログラム番号(54, Voice-Ooh)に設定して再生すれば、より音声らしい再生が可能であることは確認できた。

6. あとがき

本稿では日本語の各音素を二連の8重音以下の単純な和音で表現したMIDIコードのデータベースを構築する方法および制作ツールを提案し、それを用いて仮名テキスト入力による楽器音を用いた音声合成システムの実現性について述べた。今後は、より明瞭な音声再生が実現できるよう音素の和音構成の見直しと外国語音声への対応を含めて、音素MIDIコードのデータベースの再構築を進める予定である。



(1) 男声



(2) 女声

図7 日本語音素71音MIDIコードデータベース

男声	ア E5 D#5 F#4 G#2	イ C#3 C3 G#2 C#2	ウ G2 C#2 B1 A#1	エ A3 G#2 G2 A1	オ A3 G#3 D#3 G2
カ	D#4 F5 C#2 C5 C2 G4 B1 B1	キ A#6 D#3 D#3 D3 D3 B2 D#2 D#2	ク E3 G#3 C#3 D3 F2 C3 C#2 C#2	ケ G#3 A#3 F3 A3 A#2 G3 A1 C2	コ A#3 B3 D3 A#3 D2 G#3 C#2 G3
サ	F4 F5 E4 A4 D2 F#4 C#2 E4	シ D#3 D3 D3 C3 C3 D2 D2 C2	ス G3 G#3 C#3 E3 C#2 C#3 C2 C3	セ G#5 B3 G#3 A3 D3 G#3 A#2 F3	ソ C4 C4 F#3 A#3 B2 G3 C2 G#2
タ	E2 D#5 D#2 A4 C#2 G#4 C2 F4	チ D3 D#3 C#3 D3 C#2 C3 C2 E2	ツ G3 C3 C3 C#2 C2 C2 B1 B1	テ B3 B3 A#3 A3 E3 F3 C2 B2	ト C4 C4 B3 B3 C2 A3 A#1 B1
ナ	G#2 G5 G2 F#5 G1 A#1 F#1 G1	ニ C3 D#3 G2 D3 F#2 C3 A#1 C2	ヌ A2 G#3 G2 C#3 A#1 C3 G1 C#2	ネ G#2 A#3 G2 G#2 F#2 G2 F#1 G#1	ノ C3 C4 A#2 B3 G#2 A#2 B1 C2
ハ	C#2 D#5 C2 C#5 B1 A#4 A1 F#4	ヒ D#3 D3 D3 C3 C#2 D2 C2 C2	フ C#3 C#3 F2 C3 C#2 D2 C2 C#2	ヘ C3 F3 C2 E3 B1 D3 G#1 A1	ホ E4 F4 G#3 C#4 G3 G#3 C2 C#2
マ	A2 E5 G#2 A#4 A#1 F#4 A1 B1	ミ A2 D3 G#2 C3 G3 G#2 A1 C2	ム G#2 D3 G2 G2 B1 C#2 A#1 C2	メ G#2 A3 G2 A2 F#2 E2 G#1 G#1	モ F#3 C4 B2 G3 A#2 B2 A#1 C2
ヤ	F#3 G4 G#2 E4 A#1 B3 A1 C2		ユ B2 E3 A#2 D#3 G#2 D3 G2 C#2	ヨ C3 A3 A2 G#3 G#2 B2 G2 C#2	
ラ	A3 G5 D#3 E4 A2 C4 A1 C3	リ C3 C#3 F#2 C3 A#1 C#2 A1 C2	ル C3 D3 G#2 C#3 F#2 C3 G1 C#2	レ D3 A#3 E2 A3 G1 E3 F1 A#2	ロ G#2 B3 E2 C3 B1 B2 G#1 C2
ワ	D4 G4 G#2 F4 A1 D4 G1 A1		ヲ G#3 A3 F#3 G#3 B2 G3 A#1 C#2	ン D3 B2 C#3 D2 C#2 C#2 C2 C2	
ガ	G2 C#5 F#2 D#4 G#1 B3 G1 B1	ギ G#2 D3 C2 D#2 G#1 C#2 G1 C2	グ G#2 D#3 D2 D3 G#1 C#3 G1 C2	ゲ G2 A3 E2 G3 D#2 B2 F#1 A1	ゴ C#3 G#3 C3 G3 G1 C#2 F#1 C2
ザ	G2 E4 F#2 C4 F2 C3 F#1 C2	ジ C3 D#3 F#2 C3 G1 A#2 F1 E2	ズ F#2 D3 E2 C#3 G#1 C3 F1 C2	ゼ B2 C4 E2 A#3 D#2 A3 B1 G#3	ゾ F2 A3 E2 C3 C2 B2 B1 C2
ダ	C#4 G4 C4 C#4 C2 C3 G1 C2	チ C3 C#3 C2 C3 F#1 B2 F1 G#2	ツ D3 D#3 G2 D3 G1 C#3 F#1 C#2	テ G3 A3 G2 G3 B1 F#3 G1 B1	ト G#3 B3 G2 G3 F#2 C2 G1 G1
バ	C#4 E5 A2 G4 A1 E4 G#1 C2	ビ C#3 D3 G2 C#3 C#2 C3 G#1 B2	ブ C#3 F#3 C2 F3 A#1 D#3 G#1 C2	ベ B3 A#3 F#2 A3 G1 G2 F#1 A1	ボ G#3 C#4 G3 A3 G2 G#3 G1 G#2
パ	G#4 D#5 D#2 C#5 C#2 G#2 B1 A#1	ピ D#3 D#3 D3 D3 C#3 C3 F2 D2	プ D#3 E3 D3 D#3 C#3 D3 C2 C2	ペ C4 C4 C3 A3 G#3 A3 A#2 G#3 D2 C2	ポ C4 C4 C3 A3 D#2 G3 D2 C2

図8 音素 MIDI データベース例 (男声編)

女声	ア B5 A#5 D5 B4	イ E3 D#3 C#3 B2	ウ D#3 D3 C#3 C3	エ F4 D4 C4 B3	オ A#4 D4 C4 B3
カ	A#5 G5 D5 D#5 A#4 D#4 E4 D#3	キ F#6 F#6 F#3 F3 C3 D3	ク F#3 E3 E3 D#3 D#3 C#3	ケ D#6 E6 G3 D6 E3 E4 D#3 E3	コ A#4 C#5 A4 A4 D4 G4 D#3 D4
サ	G6 B5 G5 A#5 A#4 B4 F3 A#4	シ E3 F3 D#3 E3 C#3 D3	ス F3 E3 E3 D#3 D3 C3	セ C4 D#6 A3 A#4 E3 D#4 D3 D#3	ソ F5 D5 D#5 A4 D4 G4 C#3 D4
タ	A#4 B5 A4 G#5 E3 E5 D3 B4	チ B6 A#6 E4 A6 F#3 D#3 E3 D3	ツ C6 F3 A#5 E3 F3 D#3 E3 D3	テ A#3 D6 A3 B4 D#3 D#3 D3 E3	ト G5 E5 A#4 B4 E3 D#4 E3 D#4
ナ	C#3 B5 C3 A#5 B2 G#5 A#2 B4	ニ B3 B3 C3 D#3 B2 C3 A#2 B2	ヌ C#3 G6 C3 E3 B2 D#3 A#2 D3	ネ C3 G#6 B2 E6 A#2 A#2 A2 B2	ノ C3 B4 B2 E4 A#2 D#4 A2 B2
ハ	G5 A#5 E5 G5 B4 E5 A#4 B4	ヒ G#6 G#6 G6 F#6 D#3 F3 D3 D#3	フ G3 G#5 F#3 F3 D#3 E3 D3 D3	ヘ D#6 E6 E3 E4 D#3 D#4 D3 D4	ホ C5 C5 F4 F4 E3 E4 D#3 F3
マ	C#3 A5 C3 C5 B2 C3 A#2 B2	ミ D#3 F#6 D3 E3 C3 D#3 B2 C#3	ム B2 E3 A#2 D#3 A2 D3 G2 C#3	メ C#3 G#6 C3 E4 D#3 D#4 A#2 D4	モ B2 D5 A#2 A#4 A2 D#4 G#2 D4
ヤ	C#3 A#5 C3 G5 A#2 E5 A2 B4		ユ C#3 E3 C3 D#3 A#2 C3 A2 A2	ヨ C#3 D#5 B2 D5 A#2 A#4 A2 G4	
ラ	C3 A#5 B2 D5 A#2 B4 A2 A#4	リ G#3 F#6 D#3 E3 A#2 D3 G#2 G#2	ル A#2 D#3 A2 D3 G#2 C#3 G2 G2	レ C#4 D#6 C4 D#4 B3 D4 G#2 D3	ロ C4 D#5 B2 D4 A2 C4 G#2 B2
ワ	F5 B5 A#4 A5 C#4 G5 C#3 A#4		ヲ B4 A#4 A#3 E4 D3 D#4 B2 D4	ン C4 E3 C#3 D#3 B2 D3 A#2 B2	
ガ	G4 G6 C4 B5 B3 G5 C3 C#5	ギ C#3 D#3 A2 D3 G#2 A#2 G2 G#2	グ C#3 E3 B2 D#3 A#2 D3 G2 B2	ゲ C#3 D#6 C3 C6 B2 D#4 A#2 C#4	ゴ C4 A#4 A2 E4 G#2 D#4 G2 C4
ザ	G#4 A#5 B2 D#5 G#2 D5 G2 C#4	ジ B2 A6 G#2 G#6 F#2 F#6 F2 E3	ズ A#2 G6 A2 B5 G#2 F3 G2 D#3	ゼ G#2 D#6 G2 D#5 F#2 D#4 F2 D4	ゾ B3 D#5 A#2 D5 G2 A#4 F#2 D4
ダ	G4 D#5 F#4 C5 B2 C3 F#2 F#2	チ B2 G#6 A#2 F3 G2 E3 E2 D#3	ツ A#2 G#6 A2 B5 G#2 E3 E2 D#3	テ C4 D#6 B2 D#4 G2 G3 F#2 G2	ト C#4 D#5 C4 A#4 B3 D#4 C3 D#3
バ	C5 A#5 A2 G5 A1 E4 G2 B4	ビ C#3 E3 C3 D#3 B2 C#3 A2 B2	ブ C#3 E3 C3 D#3 B2 C#3 G#2 G#2	ベ C4 E4 A#3 D#4 C3 D4 A#2 C3	ボ A4 D#5 C#4 A#4 D#3 D#4 C3 D#3
パ	D#5 A#5 B4 G5 E3 E5 D3 D#5	ピ G3 C7 F#3 F#6 E3 E3 D#3 D#3	プ F#3 G#5 D#3 F3 D3 E3 C3 D3	ペ D6 G6 E4 D#6 E3 D#4 D#3 D#3	ポ A#4 C5 D#4 B4 D#3 A#4 D3 D#4

図9 音素 MIDI データベース例 (女声編)

アイウエオ

カキクケコ

(1) 男声

アイウエオ

カキクケコ

(2) 女声

図 1 0 音素 MIDI コードの五線譜変換例

こんにちは

(1) 音声合成結果画面 (男声)

こんにちは

(2) 同五線譜変換結果 (男声)

こんにちは

(3) 音声合成結果画面 (女声)

こんにちは

(4) 同五線譜変換結果 (女声)

図 1 1 仮名テキスト入力による音声合成例「こんにちは」と同五線譜変換例