

Windows上の音合成システム「おっきんしゃい」

小坂 直敏

NTT コミュニケーション科学基礎研究所

osaka@bri1.ntt.co.jp

概要: 本稿では、音合成、加工、演奏の機能を持つ視覚的音操作環境について述べる。システムは基本的にはウィンドウ上で動作するオフライン波形エディタである。特徴的な機能としては、正弦波モデルに基づく音加工があり、この中にモーフィングも含まれている。ひとつの音はほかの音オブジェクトと演算オブジェクトを用いて操作される。また、音オブジェクトに施された手続きの履歴が視覚的に表示されるなどの特徴がある。システムは現在作曲家、あるいは音大生に使用してもらい意見交換を始めている。これは既存の楽器の延長上に存在し、将来コンピュータ音楽の創作および演奏に重要なツールとなりえる。

Sound generation system "Otkinshi" on Windows

Naotoshi Osaka

NTT Communication Science Laboratories

Abstract: This paper describes a visual sound manipulation system which includes sound synthesis, modification and performance functions. It can run on windows and be used basically as off-line wave editor. Its distinct function is sinusoidal-model-based sound modification, and it also includes timbre morphing capabilities. A sound is manipulated using other sound objects and operation objects, and the procedure history for a sound object is visually displayed. The system is now being used and tested by some composers and music school students. It is becoming an extension of conventional acoustic musical instruments, and therefore will become an important tool for computer music creation and performance in the future.

1 はじめに

音合成とその加工、およびそれらの演奏などの音操作は、コンピュータ音楽の創造でもっとも重要な点の一つである。音合成のひとつの方向性は MAX/MSP [URL1], Kyma[URL2] などに見られるような実時間合成である。しかし、非実時間音合成も今なお重要であり、通常の音楽ワークステーションには実装可能な波形エディタがいくつかある。

特に音合成のアルゴリズムによっては信号は実時間では処理されず、結果として非実時間システムとなる。ここで提案するシステムは、*Otkinshi*(おっきんしゃい; *Oto to koega isshoni naru shisutemu*) と呼ばれ、オフラインの正弦波モデルによる音加工機能を基軸に一般的な波形エディタの機能を持つものである。すでに報告したバージョン1は NeXT 計算機上で Objective-C で書かれている [1]。今回のバージョンは C++ で書かれ、Windows95,98 および NT 上で特殊なハードウェアを必要とせず動作する。このシステムは多くの音楽家、音大生、あるいは音に興味を持つ学生などに簡便に使ってもらいたいとしたいものである。本稿ではシステムの機能を中心にして、システムの構成について述べる。

2 システムの特徴

本システムの基本的な機能は、正弦波モデルを基にした音加工であり、モーフィング機能を含む。物理モデルによるモーフィングも取り込まれている。これらの機能により豊富な音色を合成することができる。

ここでは洗練された GUI とするために、以下の点を考慮した。

1. より少ないマウス操作
2. 音と演算の両方に統一的な「オブジェクト」という概念を付与
3. 手続き履歴の視覚的表示
3. はたとえば一連のフィルタリング処理のようにシーケンシャルな手続きが表現できる。これは視覚的に完全なプログラミング環境へと発展させたい。

2.1 音合成

システムの主要機能の一つは正弦波モデルに基づいた音の加工である。正弦波モデルの分析/合成は、文献 [2] が用いられている。この枠組みで簡易な加工として、部分音の振幅や瞬時周波数のいき値、あるいは範囲、のほか、部分音の長さ、あるいは存在す

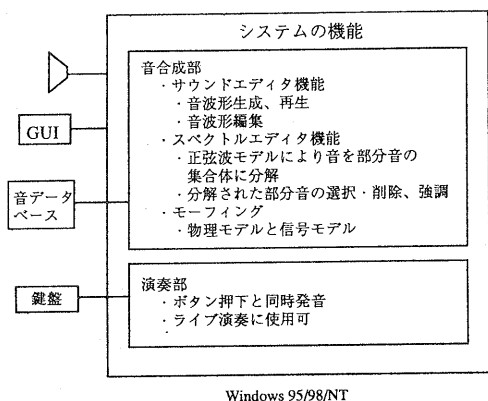


図 1: システムの構成

る周波数帯など、ある条件を満たす部分音を削除/選択することである。近い将来ビブラートの抽出とこれの付加/削除についても実装したい。

2.2 音色モーフィング

音色モーフィングは正弦波モデル表現の枠組みの中でもっとも洗練された技術の一つである。この技術の詳細は筆者の文献 [3] に述べてある。

モーフィングはモデルパラメータの補間により行われる。このアルゴリズムが解決している中心的な問題は、メンバ数の異なる二つのグループでの対応する相手の探査を最適に行うことである。このアルゴリズムを適用することにより、二つの音の対応する部分音探査を自動的に行うことができ、モーフィングが可能となる。

信号モデルと同様に物理モデルのモーフィングも実装を始めている [4]。音色モーフィングは物理パラメータを補間することにより行われる。現在は打弦、撥弦あるいは弾性体を扱うことができる。

3 システムの構成

図 1 に総合システムの構成を示す。システムは大きくわけて、音合成部と演奏部の二つの部からなる。以下は音合成部を中心に述べる。

3.1 GUI

このシステムは音オブジェクトと演算オブジェクトを統一的な概念で扱う。これらのオブジェクトは多層で再帰的である。最上のレイヤ (レイヤ 0) では、アイコン付きのボタンで表される (演算/音アイコン)。演算アイコンに触れると着目する音ファイルにその演算が実行される。音アイコンに触れると音が発する。これはそれぞれのオブジェクトにとって、最も簡単で本質的な機能である。

レイヤ 1 は、ダブルクリックにより到達する。演算オブジェクトでは詳しいパラメータ設定が行える。

図 2 は音オブジェクトの加工の過程を示したものである。図上はレイヤ 2 の音オブジェクト、一般的な音加工オブジェクトパネル (左下)、この中にあるピッチ変換オブジェクトのレイヤ 2 (右下) である。

音オブジェクトはモニタ部と演算履歴表示部からなる。モニタ部では、波形/スペクトル表示、音の受聴などからなり、履歴部では、系図のように自分のルーツからの演算履歴が表示される。

履歴表示部では、演算の履歴が音オブジェクトと演算オブジェクトを表すアイコンで表示される。この例では音ファイルの読み込み、波形の部分削除、反転、サンプリングレート変換などのシーケンシャルな手続きが示されている。

これは、視覚的なプログラミング環境の一部ともいえ、この上で編集可能である。この考えを用いて、ユーザは任意の場所で、手続きを再計算したり終了したりすることができる。MAX のパッチとこの履歴記述部との相違は、前者は手続きの定義であり、後者は音アイコン、オペレーションアイコン、あるいは制御アイコンを用いた音データの定義である点である。この手続きの履歴表示は、分岐、ジャンプ、ループなどのもういくつかのフロー制御機能を付加することによりプログラミング環境となる可能性を示している。しかし、現在はシーケンシャルなフローしか扱えず、プログラミングの観点からは非常に制約のある形しか実現できない。

3.2 スペクトル表示

スペクトルには多くの情報があり、それらを一つの方法で表示するのは不十分であり、本システムではスペクトル表示が何種類か行える。図 3 にはスペクトルが 3 種類の表示で行われている。図は原波形 (左上)、全フレームのスペクトルの 3D 表示 (左下)、スペクトログラム (右上)、正弦波モデル分析より得られる部分音の軌跡 (右下) を示す。そのほか 1 フレームのスペクトルの 2 次元表示もある。

3.3 音合成および加工の機能

一般的な音合成機能として、正弦波、鋸波、三角波、白色雑音などの基本関数、周波数変調、振幅変調、各種 FIR フィルタ、ピッチ変換、サンプリングレート変換、などがある。

3.4 演奏部

これはほかのモジュールとは独立しており、ライブコンサートでの使用も目的としている。手動演奏モード (楽器モード) では、制作された音がディスプレイ上のボタンをマウスで押すか、コンピュータキーボードを押下するだけで発音する。これらの個々のボタンは一つ一つの音ファイル、あるいは音合成部

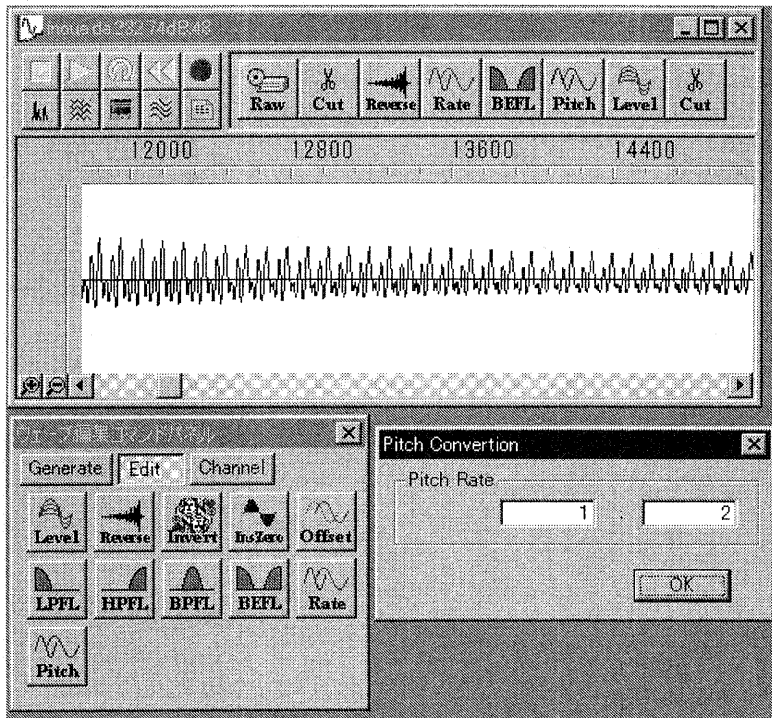


図 2: 音オブジェクトの加工過程

で作成された音オブジェクトに対応する。また、MIDI 制御も可能であり、鍵盤からの制御も可能である。

マウス、キーボードあるいは鍵盤の押下時刻と発音時刻の間には遅延はない。これはすべてのディスクファイルの音の冒頭部をメインメモリ上に格納することで実現できた。Windows の性質上、システムが個々の音オブジェクトのメモリを解放しないよう、周期的にこれらにアクセスしてやる必要がある。この機能の実現により、合成した音を他のサンプリャなどに転送する必要はなくそのままの状況で使用できる。その他の機能として、複数サウンドファイルの自動連続発音機能、個々のサウンドファイルのループ発音などがある。

4 システムのモニタと演奏履歴

本システムは国立音大にインストールされ、音楽および音の教育目的で使用されている。1999 年 1 月には NTT ICC(インター・コミュニケーション・センター)でニュースクール #6 なる企画が催され、システムが披露され、25 人ほどのユーザがシステムを体験するワークショップも行われた。

また本システムを用いたコンピュータ音楽作品も二つほど実現した。ひとつはピアノとコンピュータの作品 [P1]、また、もうひとつはオーケストラとコンピュータのための作品 [P2] である。これらはシス

テムトラブルもなく成功裏に終わった。これにより、今後の生演奏に有用な音操作ツールとなることが期待できる。

5 まとめ

音合成、加工、演奏システム「おっきんしゃい」のバージョン II について報告した。音は正弦波モデルで表現され加工することができる。この枠組みでもっとも洗練された機能はモーフィングである。また、これは一般的なウエブエディタの機能も有している。バージョン II は、音楽家、音大生、あるいは音に興味を持つ学生が使うことを目的としている。今後これらの方々に、モニタとしてソフトウェアを使用して意見をフィードバックしてもらうことにより改良を行っていく。さらにユーザインターフェース、物理モデルを改良していく。また、ビブラートとの抽出と制御、ピッチ変換あるいは音の伸縮などの加工の機能も実装していく予定である。

謝辞 この研究を行う機会を与えていただいた萩田紀博メディア情報研究部長、および研究を進めるにあたり日頃議論していただくメディア表現グループの諸氏に感謝します。

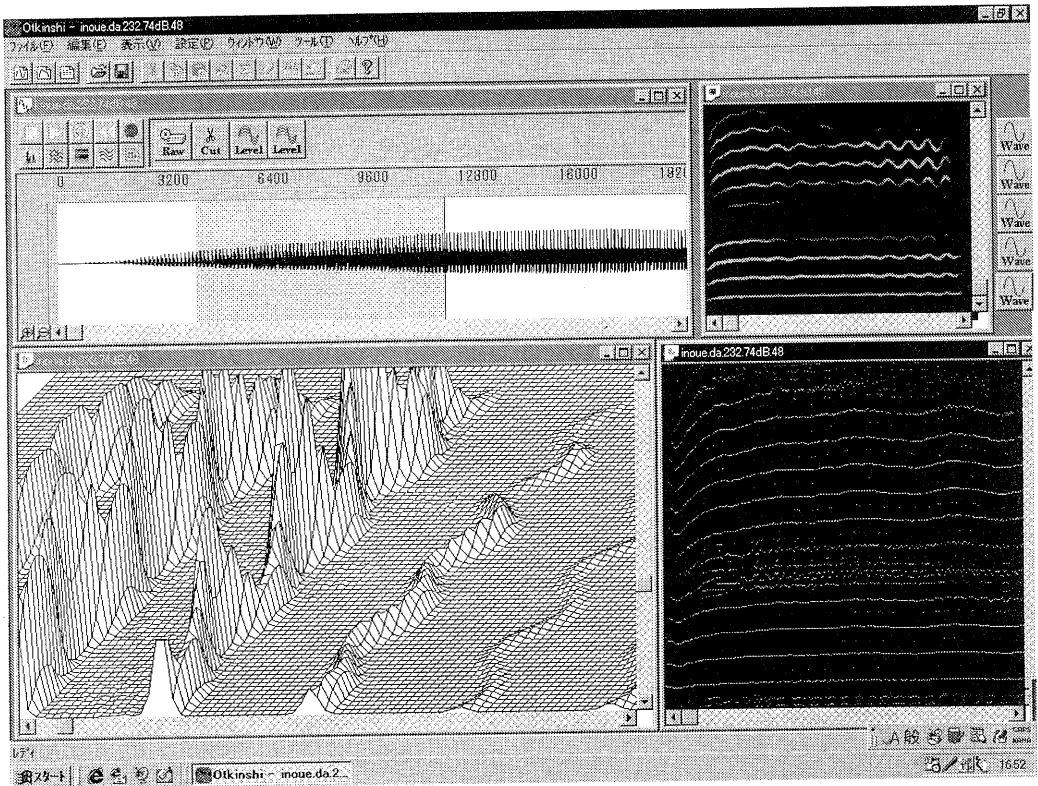


図 3: さまざまなスペクトル表示例

参考文献

- [1] Osaka, N., "Otkinshi: A sound generation and performance system," *Proc. of ICMC 92*, pp. 406-407. San Jose, California, 1992.
- [2] McAulay, Robert J., and Quatieri, Thomas F., "Speech Analysis/ Synthesis Based on a Sinusoidal Representation," *IEEE Trans. on Acoust., Speech, and Signal Processing*, vol. ASSP-34, No. 4, Aug. 1986.
- [3] Osaka, N., "Timbre interpolation of sounds using a sinusoidal model," *Proc. of ICMC95*, pp. 408-411, Banff, 1995.
- [4] 引地孝文, 小坂直敏, "Morphing of sounds of the struck strings, plucked strings, and elastic media," 信学技報, SP96-111, pp. 23-28, Feb. 1997.

参考 URL

- [URL1] <http://www.cycling74.com/>
- [URL2] <http://www.SymbolicSound.com/kyma.html>

参考演奏

- [P1] 小坂直敏, "ピアノとコンピュータ 2 台のための「音の織物」" 渋谷淑子ピアノリサイタル, お茶の水 カザルスホール, 1998.10.20.
- [P2] 小坂直敏, "オーケストラとコンピュータのための「季の崩し」" NTTCS 研 研究用音収録, 東京オペラシティリサイタルホール, 1999.3.24.