

音色記号による環境音合成のためのシステムの検討 小林 洋平 小坂 直敏

東京電機大学

〒101-8457 東京都千代田区神田錦町 2-2

E-Mail: yohei@srl.im.dendai.ac.jp, osaka@im.dendai.ac.jp

あらまし

筆者らは音メディアコンテンツ制作ツールとして、音色に記号を付与し、それらを用いた環境音合成システムを構築している。本システムを利用するにあたり、特に、GUIは膨大な量の音色記号の中から、ユーザがイメージする音色記号の素早い探査や、自由な粒度での新たな音色記号の柔軟な登録、そして簡便な音合成を行うために重要な機能である。本稿では、音色記号探査に用いるGUIのより詳細な設計、および探査方式について検討する。また、記号登録の際に有用な機能についても紹介する。

Timbre script-based sound synthesis system

Yohei KOBAYASHI, Naotoshi OSAKA

Tokyo Denki University

2-2 Kanda-nishikicho, Chiyoda-ku, Tokyo, 101-8457 Japan

E-mail: yohei.yoshida.goto@srl.im.dendai.ac.jp, osaka@im.dendai.ac.jp

Abstract: The authors are investigating a timbre notation-based environmental sound synthesis system. Because timbre notation is a concept still unfamiliar to users, the system must provide a convenient GUI for efficiently and quickly finding a target symbol and the sound it represents. In this study, several GUIs are introduced for both hierarchical structured symbols and vector space-type symbols. Several convenient functions for registration of new sound are also introduced.

1. はじめに

筆者らは、音メディアコンテンツの制作への応用を目的として、環境音を音色記号から合成するシステムの枠組みについて検討している[1]。このシステムを電子音色辞書と呼び、音楽や映像などのような、大きなサイズのコンテンツではなく、音素材のような、ひとつの音（ストリーム）あるいはその列を対象として、合成により音を提供することを主要な問題としている。

この目的では、出力音のイメージをシステムにいかんにか伝えるかが困難な問題のひとつである。音色記号の提案[2]はこれを解決するための一手段であり、音声合成や音楽合成と同様、バッチ処理による合成が可能となる。音色記号を用いた音合成システムについて、これまでに文献[1]ではその全体構想について紹介した。同システムは、バッチ処理が可能となると、音コンテンツ制作への応用のほか、

アニメーションなどの映像コンテンツへの自動効果音／環境音付加システムとして用いることができる。また、プレゼンツールなどにシステムの音色記号のデータ構造を取り込むことにより、記号を指定して音をインタラクティブに挿入することが可能となる。

また、同システムは臨場感通信において、背景音や環境音など、場面の一部に認識伝送が可能となる場合に応用が可能である。一般に、音響の研究は音声、音楽、騒音などに限定されているが、環境音を幅広く扱うことで、より豊かな通信品質が期待できる。

本稿では、ユーザのイメージに沿った探査を効率的に行うためのGUIについて、新たな提案を行う。また、音色記号の追加登録を行う際、個人や言語圏による知覚の相違や、登録時の状態により、既存の音色記号とかけ離れる場合がある。それを回避し、柔軟に音色記号の登録を行える機能について検討する。

2. システムの概要

2.1 音色記号の設計と付与

音色はその時間的粒度に応じて、様々な記述方式が存在する。粒度が小さいものでは、信号処理分野の議論となる。有限個の記号としては、ベクトル量子化がこれに相当する。また、明確に有限個の記号とは意識されないが、正弦波モデル [3] や、SMS(Spectral Modeling Synthesis)[4] など数 msec 単位の表現法が挙げられる。また、粒度の大きいものでは音韻/音素や楽音[5]、および MIDI 信号[6]が挙げられる。本研究では、音色記号は、音韻と同等の粒度である擬音語を用いる。またこのほか、信号処理レベルと音韻レベルの間を行く数十から 60msec 程度のより小さな粒度である微視的音色と、擬音語をいくつか含むような巨視的音色を定義する。擬音語はわれわれが日常使用している音色記号であり、これを拡張して用いることとする。そして、これらを元に音色を分類し、記号を付与する。表 1 に水音の分類と記号を付与した例を示す。

表 1 水音の分類と擬音語の例

区分	生成要因			音色記号	
	大	中	小	巨視的音色	擬音語
自然音	水	雨	小雨	s#	a:サー, b:シャー
			豪雨	z#	a:ザー
		川	せせらぎ	chy*	a:チヨロチヨロ
			急流	d#	a:ドー, b:ドドド

2.2 電子音色辞書編集システムの構築

2.1 で付与した音色記号をデータベースとしてネットワーク上で共有し、多くの人が音色記号の編集やそれらを組み合わせた音合成を行える電子音色辞書編集システムを構築する。図 1 にシステムのブロック図を示す。ネットワーク上からコマンド入力、GUI 操作、信号の直接入力によって電子音色辞書にアクセスし、音色記号の探査や編集を行う。また、探査された音色記号、またはあらかじめスクリプトなどで記述された音色記号列を入力とし、合成部で音合成を行う。

信号の直接入力は、図では省略したが、二つの用途があり、一つは認識を経て、音色辞書にアクセスする方法、またもう一つは合成部に直接関与する方法である。前者は将来的に検討したい。後者は、懇声音(sound hybrid)の枠組みであり、音色記号では反映できない補助的な部

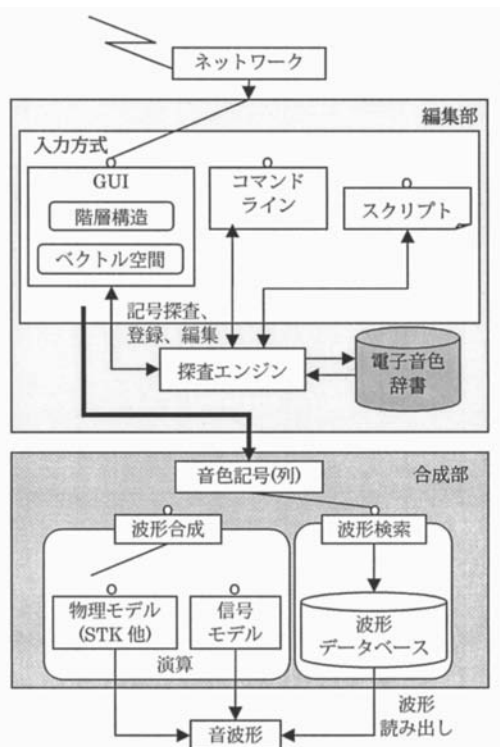


図 1. 電子音色辞書編集システムのブロック図

分を受け持つ場合と、音色記号による合成の枠組みをはずれ、信号から抽出された特徴量を直接合成部への入力とするものである。文献[7]でその一例を紹介した。

3. 探査機能の実際

本システムの探査に関連するものとして、大量の映像や音楽などのマルチメディアコンテンツの中からユーザの嗜好や目的に合った検索技術が注目されている。楽曲検索の分野では、類似度を用いて近傍の楽曲を検索する Musicream[8]や、楽曲構造の区間を抽出し、先頭へ瞬時に早送りする SmartMusikKIOSK [9]が存在する。また、楽曲に対する多様な解釈を扱うアノテーションシステム[10]や、SoundComplete[11]などは、楽曲の特徴の編集や、それらを用いた新たな楽曲の制作が可能である。しかし、環境音に対してこれらと同様の研究はなく、音データを順番に聞いて判断する、あるいはファイル名を用いた検索手法が一般的である。

本システムで扱う音色記号や音データも膨大になることが予想され、その中からユーザが望む音色を如何にして素早く探し出すかがシステムを利用していく上での検討課題である。そのため、従来の試聴による判断や、ファイル名などの文字列を使った手法では、作業効率の低下に繋がる。

一方、音色記号は新たに提案している記号であるため、ユーザにとって一般的ではなく、この意味でも目的とする音まで辿るのに困難なことが予想される。擬音語は辞書項目としては限定されたものしかないため、ほとんどの音はテキスト検索で見つからない可能性がある。また、このままではユーザにまったくなじみのない微視的音色、あるいは巨視的音色などの検索は絶望的である。

ここでは、扱うデータ構造が階層型とベクトル型の2種類あるため、それぞれの構造を視覚的に辿りながら探査する方法（階層型探査 GUI、ベクトル型探査 GUI）とより能動的にユーザが聞きたいと思う道筋を音色記号上で指定する蛇道試聴とを提案する。

3.1 階層型探査 GUI

音色記号はその粒度に応じて階層的な構造にすることができる。最下層は擬音語や微視的音色とし、その上位ではそれらを包括する抽象的な記号とする。この階層構造を GUI によって表現し、マウスを用いた全方位移動、画面の拡大、縮小を行うことで、ユーザの所望する粒度の音色を見つけ出すことができる。図 2 に GUI の拡大／縮小の画面例を示す。ユーザが「火」や「水」、「機械」など大まかな分類の探査や、巨視的音色を利用したいときは、階層構造を縮小し広域表示する。その際、各階層の代表となる記号を表示する。そして、「風音」の下位にある「ビュー」など、より細かい分類や階層を調べたいとき、あるいは厳密に音色を指定したいときは拡大し、擬音語や微視的音色を表示することで、任意の粒度で所望の音色を見つけ出すことができる。

3.2 ベクトル型探査 GUI

階層的な構造とは別に、音色を距離空間上で表現することにより、Google Earth と同様の近傍探査、移動などが可能である。そこで、音色に関する特徴を数値として複数設け、それらを要素とするベクトル空間を考える。図 3 にあ

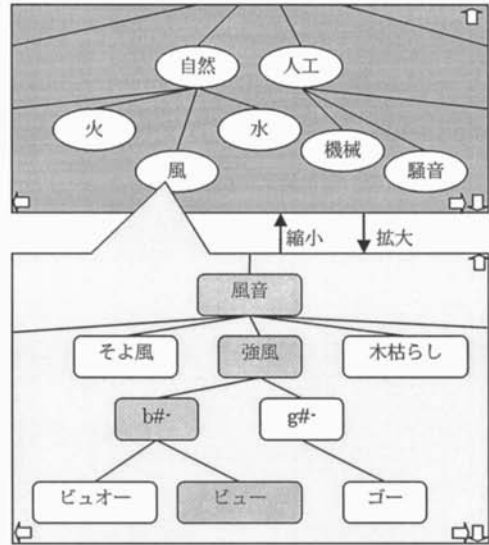


図 2. 階層型 GUI の拡大／縮小の画面例

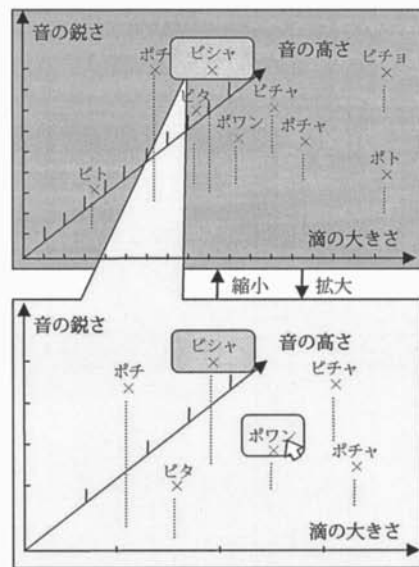


図 3. ベクトル型 GUI の拡大／縮小の画面例

る水滴音を選択したときの GUI の拡大／縮小の画面例を示す。ベクトルの各軸に要因を任意に割り当て、現在選択している音色を中心に、他の音色を表示している。特徴が近い音色を探査するときは拡大し、逆に離れた音色を探査するときは縮小させて表示させる。各軸に割り当

てられた要因は自由に変更でき、その都度候補となる音色を自動で変更、再提示する。ユーザは、現在選択している音色が所望の音色と異なるとき、要因を参考に別の音色を選択することで、その音色を中心としてベクトル空間を更新する。これを繰り返し、所望する音色へと到達できる。

3.3 階層型探索とベクトル型探索

階層型探索 GUI とベクトル型探索 GUI とは、どのような関係にあるか、という問題が生ずる。すなわち、任意の音はどちらの型でも探索できるのかどうか、という問題である。これは音色記号のデータ構造そのものに依存する。現在は、音の生成に基づいた分類を階層的に行っている。最上位から下層にいくに連れ、細分化されるが、あるレベル以下は実際の音色記号の階層となる。すなわち、巨視的音色、擬音語、微視的音色などの順に下位の音色記号の階層となっていく。ベクトル型の構造は、同一の複数要因で音を評価できるものをフラットにくくることが基本的な考え方である。例えば楽音では音の印象に関する基本的な心理要因については、普遍的な要因を求めることが大きな研究テーマである。

現在の評価要因群をいくつか設け、一つの巨視的音色の粒度の階層以下はそこに属するベクトル型の評価要因は同一であるものとする。例えば、図3で示すベクトル形構造は、巨視的音色 p# 以下の階層では、みな同一の構造を有するものとする。

従って、広角から拡大まで音色のダイナミックレンジの大小に関係なく探索するには階層型 GUI による探索が便利である。ベクトル型探索は任意の広範な音をベクトル空間上で必ずしも比較することができず、相対的に音色距離の近いところでの比較に適している。しかし、バイオリンのピチカートと水滴の音とを比較しながら音を探査、選択する場合を考えると、楽音と同じ要因を水滴に設定しているか、というシステム側の音色データベースに依存し、必ずしも同一でない場合もある。これは現在の階層型分類が生成に基づいているため、こちらでの音の距離が遠く、階層型の構造に基づいてベクトル型データを決定していることにより生ずる問題である。

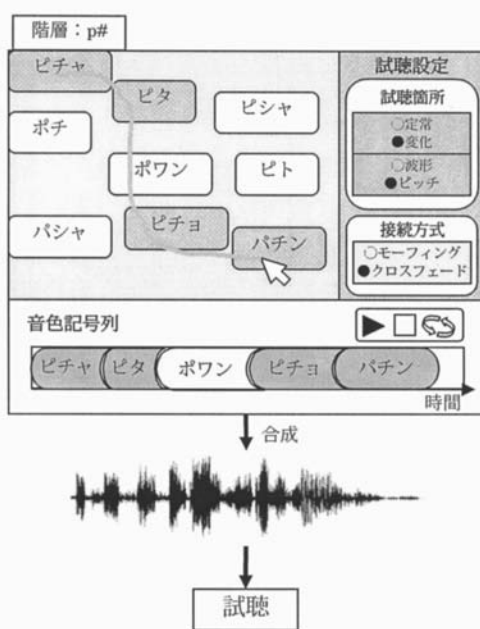


図4. 蛇道試聴のブロック図

3.4 蛇道試聴 GUI

3.1、3.2 はデータ構造に基づく音色の探索であるが、音を扱っている以上、実際に聞いて判断するという場合も多く存在する。そこで、ある階層、またはあるベクトル空間において、音色記号群をマウスカーソルなどでなぞる、囲うなどの操作を行うだけで対応する音色を連続して試聴することができる蛇道試聴の機能を提案する。図4に動作のブロック図を示す。ここでは、音色記号「p#」内の音色をランダムに並べ、それらをマウスカーソルで自由になぞったときの動作を表している。カーソルの軌跡と重なる音色記号に対応した音データを順に接続し、再生する。また、これらを音色記号列として時間軸上で表現することで、再生順、再生位置を確認できる。これにより、ユーザのイメージと合う音色を記号という視覚情報だけでなく、実際に聞くことでより正確に見つけ出すことができる。また、再生も単純な連続再生だけではなく、クロスフェードやモーフィングを用いた滑らかな合成、音データの波形またはピッチにおける定常部や変化部の部分再生など、細かい指定を行うことで、より素早い探索が可能となる。

4. 登録補助機能

本システムでは、利用できる音色記号の幅を広げるために、探査だけでなく新たな音色と記号の登録する機能を設ける。しかし、個人や言語圏の違いにより、知覚的に近い音色でも記号がかけ離れている場合や、逆に近い記号でも音色が知覚的に大きく異なるという場合がある。また、同一ユーザであっても登録するたびに投入する音色が異なりうる、という再現性の問題もある。さらに、微視的音色を登録する際、時間的粒度が小さく、それ単体では知覚できない場合がある。以上のように、音色記号の登録時には、これを定義しにくい潜在的要因が数多くある。そこで、これらの問題を解決するための登録時の補助機能について検討した。

4.1 類似音色記号による確認機能

ユーザがある音色とそれに付与する記号を登録する際、システムは一旦電子音色辞書内のデータを自動探査し、同じ、あるいは近い音色記号を提示する。ユーザは実際にその音色を聞くことで、登録する音色記号が最適かどうかを判断し、必要に応じて修正を行うことが可能となる。図5にある波形データと「ドン」という記号を入力し、登録する際のブロック図を示す。

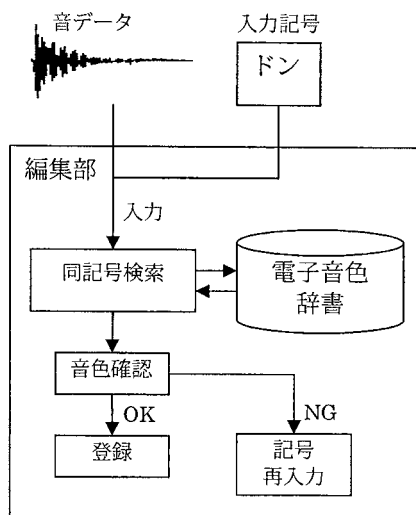


図5.類似記号による確認機能のブロック図

4.2 音色と記号の自動対応付け機能

環境音は音色記号を複数組み合わせさせた擬音語の粒度で表されるものが多い。これらを微視的音色の粒度まで細分化して一つひとつ登録することは知覚的に困難であり、また時間もかかる。そこで、擬音語などの大きな粒度の音色と、知覚した音色記号列を入力し、システム内部で音色の細分化と、記号との自動対応付けを行う機能を検討している。図6に水滴音と擬音語を入力したときのブロック図を示す。さらに、細分化され、記号が付与された音色に対して4.1の 確認機能を用いることで、より最適な音色記号への修正が可能となる。これは音声のラベリングデータと音声信号が与えられた場合の自動セグメンテーションの技術と関連する。

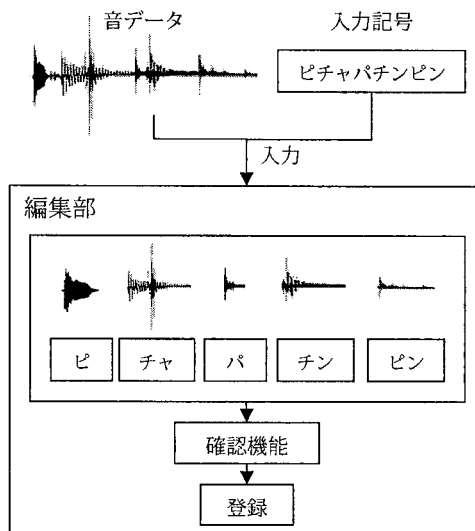


図6.自動対応付け機能のブロック図

5. おわりに

本報告では、音色記号を用いた電子音色辞書編集システムの GUI と登録時の補助機能について検討した。特に、本研究の目的のひとつである、より素早く音色記号を見つけ出すために、

階層型、ベクトル型とともに、蛇道試聴 GUI を新たに提案した。蛇道試聴 GUI は音色の階層構造表示上画面でユーザが自由に音を選択し、これらがまとめて再生されることで、システムの素早い反応に対する要求条件を満たそうとするものである。

今後は、音色の分類や記号付与を進めるためのオンライン上での共有を目指す。また、より

具体的な探査 GUI の構築と、実際の音合成に必要な GUI の検討も行う。

参考文献

- [1] 小林, 吉田, 後藤, 小坂, “音色記号による音合成のための電子音色辞書編集システムの検討”, 情報処理学会研究報告, 2007-MUS-69, pp.19-24, 2007
- [2] 小坂 直敏, “楽曲制作のための音色理論の構築にむけて”, 日本音響学会講演論文誌, 3-7-15, 2005.
- [3] McAulay R. J. McAulay and T. F. Quatieri, “Speech Analysis/ Synthesis Based on a Sinusoidal Representation,” IEEE Trans. on Acoust., Speech, and Signal Processing, vol. ASSP-34, No. 4, Aug. 1986.
- [4] Serra, X. and J. Smith, “Spectral modeling Synthesis: A Sound Analysis/Synthesis System based on a Deterministic plus Stochastic Decomposition.” Computer Music Journal 14(4): 12-24, 1990.
- [5] Rodrigo Segnini, “Score phonetization and speech derived composition,” Proc. of ISMA (International Symposium on Musical Acoustics), 2-S2-13, Nara April 2004.S
- [6] Modegi, T., “Development of MIDI Encoder Tool ‘Auto-F’ for General Time-Based Electric Signals,” ICAD2002 (The Eighth International Conference on Auditory Display), pp158-163, Kyoto, Japan, 2002.
- [7] 吉田誠一, 小坂直敏「音響信号の直接入力による概念からの音合成」情報処理学会第 69 回全国大会, 2-225, 2007.
- [8] 後藤 孝行, 後藤 真孝, “Musicream: 楽曲を流してくっつけて並べることのできる新たな音楽再生インタフェース”, WISS 2004 論文集, pp.53-58, 2004.
- [9] 後藤: “SmartMusicKIOSK: サビ出し機能付き音楽試聴機”, 情報処理学会論文集, pp.9-16, 2003.
- [10] 梶, “楽曲に対する多様な解釈を扱うアノテーションシステム”, IPSJ 論文誌 Vol.46 No.4, 2005
- [11] 平田, 松田, “音楽エンタテインメントソフトウェア SoundComplete の現状報告”, IPSJ 研究報告, 2003-MUS-52