

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

小林洋平[†] 小坂直敏[‡]東京電機大学^{†‡}

1. はじめに

現在、映画やゲームをはじめとするマルチメディアコンテンツにおいて、音楽、音声、効果音が様々な場面で使用され、これらに用いられる音合成技術やシステムは数多く存在する。この中で、音楽と音声はスクリプトから自動合成が行えるが、効果音は試行錯誤により合成されるのみで、自動合成が行える枠組みはなく、合成方式としては不便である。また、音楽、音声と同様、効果音もその表現、記録、伝達の目的で時間情報を有す音色系列として記述する必要がある。

以上を考慮し、本研究では音色を記号として定義[1]し、それらを組み合わせた記号列を用いることで自動合成を可能とする。特に効果音制作時には、制作者の意図に沿った音を合成する柔軟なシステムの構築を目的とする。

2. 現在の音合成システム

映画やゲームなどで使用される音コンテンツは音楽、音声及び効果音の3つに分けられる。以下に現在の合成方式の動作概要を述べる。

2.1 音楽合成

一般にシーケンサと呼ばれる音楽合成システムは、楽譜を入力として合成する。システム内部では楽譜情報を元に音の高さを表す記号の列である音高列と、音の長さを表す音長列の二つを根幹とし、音の強弱などの演奏表情の情報を加え合成する。図1に動作の概略図を示す。また、MIDI信号を利用し、異なる電子楽器間で記号列、及び演奏表情の情報通信を行った合成も可能である。

2.2 音声合成

音声合成も多く多种方式が存在するが、その一つとしてテキストデータを基に音声データベースから音素片を取り出し、接続する素片編集型合成が主流である[2]。図2にブロック図を示す。入力された原文は言語解析部で読みとアクセントが付与され、韻律計算部で韻律モデルを用いて抑揚やリズムの計算を行い、発音記号列のほか、ピッチ列、音長列を算出する。そして素片接続部でそれらに基づき音素片を選択、接続し合成する。

2.3 効果音の合成

一方、これらに対し効果音の合成は、収録された音データ(サンプル音またはモデル音)にユーザが対話的にエフェクトを加えて音合成を行う。先の2例のように記号系列に基づく自動合成はなく、これまでは試行錯誤による合成手法が一般的である。

本研究では、効果音や新しい音楽に対しても音楽/音声と同様に記号系列や時間長列などを入力とし、自動合成するシステムを構築する。これにより、時間的マルチメディアのトラックに、これらを用いてテキストとして記述でき、その結果、スクリプトとして扱え、表現形式として便利になる。さらに、それを保存し、異なる環境や他者による合成の再現も可能になる。

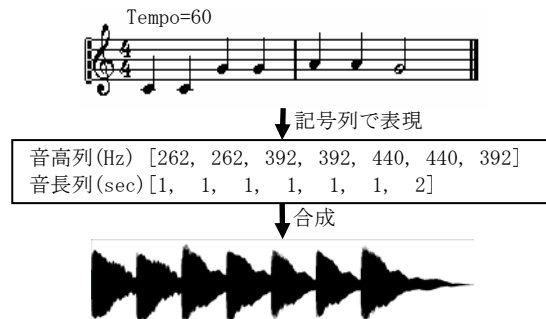


図1. 音楽合成の動作概略図

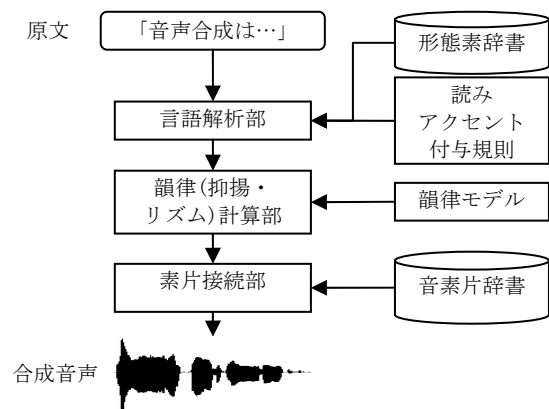


図2. 素片編集型合成のブロック図

表1 水音の音色記号と擬音語の例

区分	生成要因	音色記号	音色記号/擬音語
A	滴	p#	a:ピチャ, b:ピシャ, c:ピタ d:ポチャ, e:ポチ, f:ポワン
B	バブル(小)	s#	a:ショワー, b:シャー, c:ジャー
	バブル(大)	k*	d:コポ, e:ゴボ
		b#, d#	f:ブク, g:ドク

3. システムの概要

3.1 音色記号の定義

対象とする音を電子音や自然音などの生成状況、孤立発声や連続発声の発声状況、そして知覚に基づいて分類し、記号や擬音語を定義する。表1に水音の一部の音色記号と擬音語の例を示す。

3.2 電子音色辞書編集システムの構築

3.1で対象とする音は数多く存在する。また、音の聞こえ方は個人や言語圏などで異なり、一意でない。さらに、定義された音色記号も膨大になる。そのため、音色

A study on electronic timbre dictionary editing system for sound synthesis by timbre script.

[†] Yohei KOBAYASHI (Tokyo Denki University)

[‡] Naotoshi OKSAKA (Tokyo Denki University)

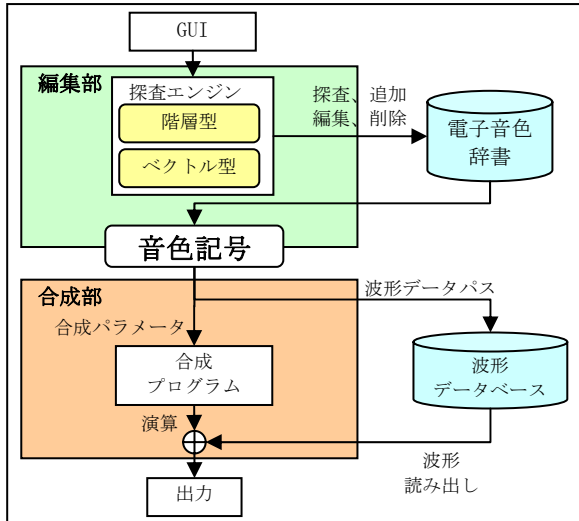


図3. 電子音色辞書編集システムのブロック図

記号の定義を個人、または数人で完成させることは困難である。そこで、wikipedia と同様に、音色記号のデータベースを共有し、多くの人が参加できる電子音色辞書編集システムを構築する。図3にシステムのブロック図を示す。

システムを利用する際、膨大かつ不慣れな音色記号から制作者の意図する音色記号を素早く見出すためには、の検索を容易にする探索機能が必要である。また、記号の追加定義や対応する音サンプルや合成プログラムの提示、特徴を要因として付加するなどの編集機能も必要である。現在は、検索に有用な2つのGUIを検討している。

3.2.1 階層型 GUI

音色記号は階層的な構造にすることができる。すなわち、最下層では擬音語とし、その上位ではそれらを含む抽象的な記号となる。制作者は、どのレベルでも記号を付与でき、この構造をGUIによって表現する。図4にある階層を辿った場合のGUIの概略図を示す。制作者は合成したい音の種類や音色記号、またその粒度によって自由に各階層を行き来し、所望の音色を見付けられる。

3.2.2 ベクトル型 GUI

3.2.1の階層的な構造とは別に、各音色が距離空間上で表現できると、近傍探索、移動など、Google Earthと同様のGUI探索が可能である。そこで、音色の階層構造とは別に、各音色に関する知覚的な複数要因を設け、各要因の数値を要素とするベクトル空間を考える。図5に近傍のある音を選んだときのGUIの概略図を示す。制作者は選択している音が所望の音と異なる時、各要因を参考に所望の音色へ素早く到達できる。

3.3 音合成システムの構築

音色空間から選択された音色記号を時系列に並べ、音合成を行うシステムを構築する。図6に音色記号を用いた効果音のスキプトの例を示す。音色記号の時間長やピッチなど、合成の際に必要なパラメータの編集機能を備え、制作者の意図に沿った合成を可能にする。また、映像や楽曲データの読み込み、再生の機能を実装することで、マルチメディアコンテンツの中での効果音の意義が確認できる。

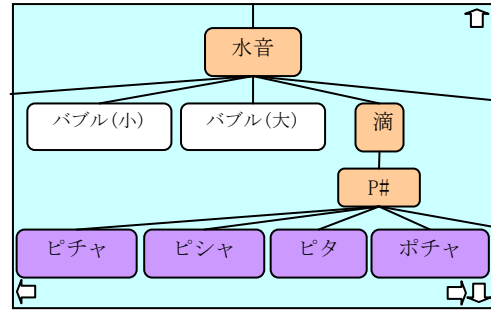


図4. 階層型 GUI の概略図

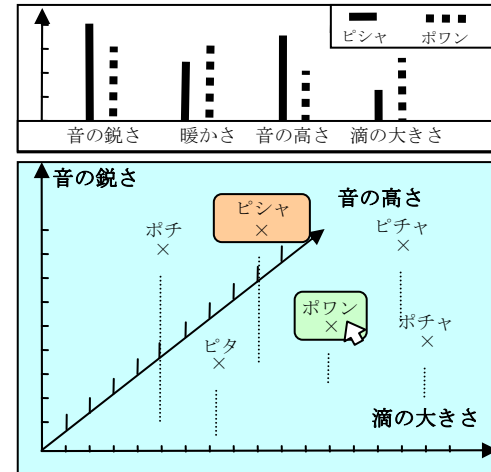


図5. ベクトル型 GUI の概略図

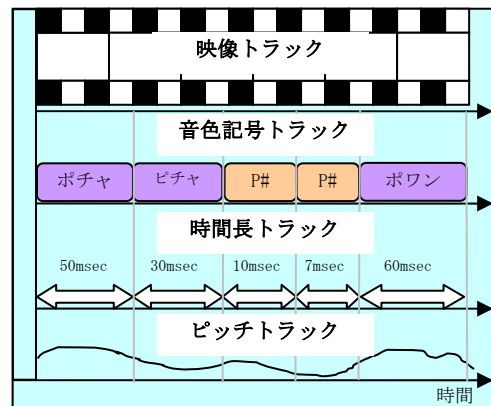


図6. 音色記号を用いたスキプトの例

4. おわりに

音楽合成、及び音声合成と同様に、記号列による電子音色辞書編集システムと、それに用いる音色記号について述べた。また、膨大な音色記号を扱うための分類方式として階層型とベクトル型のGUI方式を提案した。

今後はより多くの音色記号の定義、共有ネットワーク及び具体的なGUIの構築を進める。

参考文献

- [1] 小坂 直敏, “楽曲制作のための音色理論の構築にむけて”, 平成 17 年春季音響学会講演論文集, 3-7-15 (2005)
- [2] Akemi Iida, Nick Campbell, “Speech Database Design for a concatenative Text-to-Speech Synthesis System for Individuals with Communication Disorders”, Int'l Journal of Speech Technology, Vol.6, Issue 4, pp.379-392, (2003)