

音楽理論に基づく音声合成による感情表出手法

千葉央輝† 樽松理樹† 羽倉淳† 藤田ハミド†

岩手県立大学ソフトウェア情報学部†

1. はじめに

音声は、大きく言語情報と非言語情報から構成されている。従来の音声合成は、言語情報を中心に行われてきたが、音声合成の利用範囲の拡大とともに非言語情報にも着目されるようになってきた。その結果、感情表現を試みる音声合成手法の研究が活発化している^[1]。しかし、評価実験などにおける人の認識精度は、実験参加者や発話内容などの実験条件に依存する点が多い。よりの確な合成音声による感情表出を行うためには、更なる改善が必要となっている。

一方、音楽と感情の関係に関しては、いくつかの知見が得られている。それらを利用したシステム開発^[2]も進められている。

以上のような背景から、本稿では、音楽と感情の関係を音声合成に利用することを考え、音楽理論に基づいた新たな音声合成手法を提案する。本手法は従来手法と異なり、理論を活用することから、学習データを必要とする従来手法よりも開発効率が向上し、より実用的な手法となることが期待できる。

2. 音楽理論に基づく音声合成手法

本章では、音楽理論に基づく音声合成手法について説明する。ただし、誌面の都合上、音楽理論については簡易的な説明のみとする。

入力としては、表出を試みる感情を付与した文章を与える。ここで表出を試みる感情としては、Ekman らが提唱した怒り、喜び、悲しみ、驚き、恐れ、嫌悪の基本 6 表情^[3]を用いる。

はじめに、読み上げる文に対し、アクセントやイントネーションに基づき、各音の高さを付与する。これは日本語のアクセントが音の高さで表現されること^[4]に基づいている。本稿ではこれを基本旋律と呼ぶ。本手法は、この基本旋律を音楽理論に基づき変化させることで感情表出を試みる。

表 1: 感情とコードの対応表

感情	悲しみ	喜び	恐れ	驚き	怒り	嫌悪
コード	C _m B _m E _m	D G A	B _b _m E _m	B _b _m G _m G	F E	F _m

が 8 未満の場合、前の文節と合わせて一つの文節とする。これは、文節に含まれる音節数が少ない場合、基本旋律の変更点が少ないことから感情表出が困難になることを避けるためである。

次に、感情とコードの対応表^[5](表 1)を元に表出する感情に対応したコードを取り出す。コードとは、高さの異なる複数の音の組であり、表 1 における C_m や D が一つのコードである。さらにコードに含まれる音から、コードに対応するスケール(音の高さの組)を、音楽理論^[6]に基づき決定する。今回は、「イオニアン」「フジリアン」「ミクソリディアン」「ロクリアン」「ドリアン」「リディアン」「エオリアン」の 7 スケールを利用する。また一感情に対するスケールはできるだけ統一する。そのために、一つのコードに対応するスケールが複数ある場合は、一つのスケールを選択する。ただし、一つの感情に含まれるコードすべてに一つのスケールが適用できない場合は、複数のスケールを用いる。

次に各文節の基本旋律に含まれる音のうち、表出する感情から決定したスケールに含まれない音を、スケール中のもっとも近い高さの音に変更する。もっとも近い高さの音が複数あった場合、高い音を選択する。この処理により、合成音声のすべての音は、表出する感情で決定したスケールを構成する音となる。すなわち、楽曲における感情表出に適した音の高さを音声合成に適用することとなる。

最後にした文節ごとの旋律を結合する。このとき不自然な発話にならないように結合部分の音の高さをそろえるなどの調整を行う。

以上によって得た旋律にそって音声合成を行う。なお本手法では、音の高さのみを変化させており、速度や長さについては変化させていない。

3. 評価実験

3.1. 概要

本手法の妥当性を評価するために次に示す聴取実験を行った。実験参加者は生成した合成音声を聴き、各感情に対し、「感じた」「感じない」

次に文を文節に区切る。このとき文節の音節数

“A Study of Emotional Speech Synthesis Based on Music Theory”

†Hioki Chiba, Masaki Kurematsu {kure@iwate-pu.ac.jp}, Jun Hakura and Hamido Fujita

†Faculty of Software and Information Science, Iwate Prefectural University

「どちらともいえない」のいずれかを回答する。これらの回答結果とその合成音声で表出を試みた感情とを比較することで本手法の有用性を評価する。なお今回の実験には、男女 8 名ずつ合計 16 名が参加し、発話内容としては、内容から感情が推定しにくい天気予報を用いた。

また合成音声作成には、クリプトン社の歌声合成ソフト“巡音ルカ”^[7]を用い、悲しみ、恐れ、嫌悪に対しては 4 通り、喜び、怒りには 2 通り、驚きには 8 通りの計 24 通りのスケールの異なる合成音声を作成した。なおコードとスケールの対応数に差があるため、感情毎の合成音声数は異なる。

3. 2. 実験結果

実験結果を表 2 に示す。表 2 における一致率とは、それぞれの合成音声に対し、表出を試みた感情に対し「感じた」と答えた人数の割合である。全体として、30%台が大半をしめ、最高でも 56%にとどまった。また、同じ感情でもスケールによって一致率に差がみられた。なお表中の太字は各感情における一致率の最大値を示す。

表 2 : 各感情の一致率

悲しみ	イリアン	ドリアン	フリリアン	ロクリアン
	37.50%	31.25%	31.25%	31.25%
喜び	アイニアン	リテリアン		
	6.25%	50.00%		
恐れ	イリアン	ドリアン	フリリアン	ロクリアン
	31.25%	18.75%	18.75%	37.50%
驚き	アイニアン &イリアン	アイニアン &ドリアン	アイニアン &フリリアン	アイニアン &ロクリアン
	12.50%	25.00%	25.00%	31.25%
	リテリアン &イリアン	リテリアン &ドリアン	リテリアン &フリリアン	リテリアン &ロクリアン
	12.50%	25.00%	43.75%	18.75%
怒り	アイニアン	リテリアン		
	25.00%	37.50%		
嫌悪	イリアン	ドリアン	フリリアン	ロクリアン
	37.50%	56.25%	43.75%	37.50%

3. 3. 考察

一致率は 30%程度と低く、実験参加者毎で見た場合でも、最高で 54%、平均約 30%であった。この点から見れば、本手法の感情表出能力は低い。しかし、音の高さを変化させたのみで約 30%の一致率を得られたことから、音楽理論を合成音声に利用すること、本手法のアプローチが有用である可能性は示せたと考える。

また同じコードでも、利用するスケールによっ

て一致率が異なる。ロクリアン、リディアン、エオリアンの一一致率は 30%以上であるが、アイオニアンの一一致率は最高でも 25%と低い。これはコードに含まれる音程(2音の高さの差)の違いが影響を与えていると考えられる。アイオニアンは一定であるが、ロクリアンらは一定ではない。この変化の差が回答に影響していると考えられる。

実験の結果、本手法による感情表出能力の可能性は示せた。今後は、今回の実験結果の分析と反映、感情表出能力向上に効果があると考えられるコード進行やテンポ、強弱など現在実現していない音楽理論にかかる項目の追加、これらの有用性を確認するための評価実験、従来手法との比較を行う予定である。

4. おわりに

本稿では、音楽理論に基づき、音の高さ、言い換えればアクセントを変化させることで感情表出を試みる合成音声手法を提案した。実験の結果、システムと実験参加者との一致率は高くはなかったが有用である可能性を示すことができた。今後の課題としては、実験結果の分析と反映、音楽理論に基づく機能の追加、従来手法との比較を含む評価実験の実施などが挙げられる。

謝辞

本研究の一部は、科学研究補助費・基盤研究 B (課題番号: 20300078) および公立大学法人岩手県立大学学術研究費・連携研究の助成を受けています。また、本研究の評価実験に参加していただきました方に感謝の意を示します。

参考文献

- [1] M. Schröder: "Emotional Speech Synthesis: A Review", Proc. Eurospeech 2001, Vol. 1, pp. 561-564, 2001.
- [2] 三浦: "「怒り、喜び、悲しみ…」MIDI 楽曲に感情表現を付加するシステム「MOR2ART (モルツアート)」", DTM MAGAZINE, Vol. 178, pp. 98-99, 2009.
- [3] P. Ekman, W. V. Friesen 著, 工藤訳: "表情分析入門", 誠信書房, 1987.
- [4] 松崎, 河野: "よくわかる音声", アルク, 1998.
- [5] Marc-Antoine Charpentier: "Rules of Composition", 1642.
- [6] 北川: "ポピュラー音楽理論", リットーミュージック, 2004.
- [7] クリプトン社: "巡音ルカ", <http://www.crypton.co.jp/mp/pages/prod/vocaloid/cv03.jsp>