

歌声合成システムNNSVSを用いた斉唱の自然性に関する要因調査

西澤 佳飛^{1,a)} 山本 龍一¹ 戸田 智基^{2,b)}

受付日 xxxx年0月xx日, 採録日 xxxx年0月xx日

概要: 斉唱とは複数人が同一の旋律を歌う歌唱方式であり、複数の歌声が調和しあうことで独唱には無い深みのある歌声が生成される。歌声情報処理が発展する中、斉唱に対する情報処理の実現も大いに期待されるが、その第一歩として、斉唱が持つ音響的特徴を明らかにすることが重要である。本研究では、音響的特徴として揺らぎに着目し、各種音響特徴量の揺らぎを制御した合成歌声を用いることで、斉唱の自然性に影響を与える要因を調査する。各種音響特徴量を個別に操作可能な歌声合成手法として NNSVS[1] を活用し、発声タイミング、基本周波数パターン、スペクトル包絡特徴量系列における揺らぎを操作した歌声合成を可能とすることで、所望の音響的特徴のみの揺らぎを持つ疑似斉唱の生成を実現する。生成した疑似斉唱に対する聴取実験の結果から、斉唱において重要な音響的特徴を明らかにする。

キーワード: 情報処理学会論文誌ジャーナル, L^AT_EX, スタイルファイル, べからず集

Investigation of factors related to the naturalness of unison using the singing voice synthesis system NNSVS

KAITO NISHIZAWA^{1,a)} RYUICHI YAMAMOTO¹ TOMOKI TODA^{2,b)}

Received: xx xx, xxxx, Accepted: xx xx, xxxx

Abstract: Unison is a singing method in which multiple singers sing the same melody, and the harmony of multiple voices produces a deep singing voice that is not found in solo singing. As the information processing of singing voice has been developed, it is highly expected to realize information processing for unison, but as a first step, it is important to clarify the acoustic characteristics of unison. In this study, we focus on fluctuation as an acoustic feature and investigate the factors that affect the naturalness of unison by using synthetic singing voices in which the fluctuation of various acoustic features is controlled. By utilizing NNSVS as a singing voice synthesis method that can manipulate various acoustic features individually, and by enabling singing voice synthesis that manipulates fluctuations in vocal timing, fundamental frequency patterns, and spectral envelope feature series, we can generate pseudo-singing voices with fluctuations in only desired acoustic features. The results of listening experiments on the generated pseudo-songs will be used to clarify the important acoustic features in unison.

Keywords: IPSJ Journal, L^AT_EX, style files, “Dos and Don’ts” list

1. はじめに

斉唱とは複数人が同一の旋律を歌う歌唱方式であり、複数の歌声が調和しあうことで独唱には無い深みのある

歌声が生成される。歌声情報処理が発展する中、合唱分離 [2], [3], [4] などの手法が現れている。斉唱に対する情報処理の実現も大いに期待されるが、その第一歩として、斉唱が持つ音響的特徴を明らかにすることが重要である。本研究では、音響的特徴として揺らぎに着目し、各種音響特徴量の揺らぎを制御した合成歌声を用いることで、斉唱の自然性に影響を与える要因を調査する。歌声合成手法とし

¹ 名古屋大学 情報学研究科

² 名古屋大学 情報基盤センター

^{a)} joh.taro@ipsj.or.jp

^{b)} gakkai.jiro@ipsj.or.jp

てNNSVSを活用し、発声タイミング、基本周波数パターン、スペクトル包絡特徴量系列における揺らぎを操作した歌声合成を可能とすることで、所望の音響的特徴のみの揺らぎを持つ疑似斉唱の生成を実現する。生成した疑似斉唱に対する聴取実験の結果から、斉唱において重要な音響的特徴を明らかにする。

2. 関連研究

Goverらは、合唱分離についてWave-U-Netに基づく楽譜情報分離モデルを提案し、MuseScore General SoundFontでMIDIファイルから347曲の(合成)Bach Choraleを用いた実験が行われた。このモデルは、SoundFontSynthesisデータセットでは良好な性能を示したが、実際の合唱曲データセットでは不十分であった。[2]

また、Petermannらは基本周波数輪郭を条件とすることで分離性能を最適化する条件付きSpec-U-Netを提案された。しかし、合唱音楽のデータセットがないため、評価は総時間7分の3曲のみで行われた。[3]

Chenらは合唱曲の音源分離タスクに合成されたトレーニングデータを使用するために合唱音楽データを合成する自動パイプラインを提供した。実験で合成された合唱データで分離を行った際、周波数領域モデルのSpec-U-Netが最も良い分離性能を示した。そして、実際の合唱音楽データセットで分離を行った際、合成データがモデルの性能を向上させるのに十分な品質であることを実証した。[4]

3. NNSVS

NNSVSとは山本龍一氏によって制作されたオープンソースの歌声生成ツールである。[1]

4. 提案手法

5. 実験

5.1 揺らぎの組み合わせによる疑似斉唱の自然性

5.2 疑似斉唱の歌唱人数による影響

6. まとめ

謝辞

参考文献

- [1] R. Yamamoto, R. Yoneyama, and T. Toda. "NNSVS: A Neural Network-Based Singing Voice Synthesis Toolkit" in *ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2023, pp. 1-5
- [2] M. Gover and P. Depalle. "Score-informed source separation of choral music" in *Proceedings of the 21th International Society for Music Information Retrieval Conference, ISMIR 2020*.
- [3] D. Petermann, P. Chandna, H. Cuesta, J. Bonada, and E. Gómez. "Deep learning based source separation applied to

choir ensembles," in *Proceedings of the 21th International Society for Music Information Retrieval Conference, ISMIR 2020*.

- [4] K. Chen. "Improving Choral Music Separation through Expressive Synthesized Data from Sampled Instruments" (2022) in *Proceedings of the 23th International Society for Music Information Retrieval Conference, ISMIR 2022*.

情報 太郎 (正会員)

1970年生。1992年情報処理大学理学部情報科学科卒業。1994年同大学大学院修士課程修了。同年情報処理学会入社。オンライン出版の研究に従事。電子情報通信学会, IEEE, ACM 各会員。本会シニア会員。

処理 花子

1960年生。1982年情報処理大学理学部情報科学科卒業。1984年同大学大学院修士課程修了。1987年同博士課程修了。理学博士。1987年情報処理大学助手。1992年架空大学助教授。1997年同大教授。オンライン出版の研究に従事。2010年情報処理記念賞受賞。電子情報通信学会, IEEE, IEEE-CS, ACM 各会員。

学会 次郎 (名誉会員)

1950年生。1974年架空大学大学院修士課程修了。1987年同博士課程修了。工学博士。1977年架空大学助手。1992年情報処理大学助教授。1987年同大教授。2000年から情報処理学会顧問。オンライン出版の研究に従事。2010年情報処理記念賞受賞。情報処理学会理事。電子情報通信学会, IEEE, IEEE-CS, ACM 各会員。本会終身会員。