

進捗報告

1 今週やったこと

- 先行研究 (藤井さん) の確認

2 先行研究について

絵描き歌自動生成システムの成果についてまとめる.

- スケッチの類似性を評価する指標を提案.
- 手描きスケッチをストローク (1 筆) に分解して, 類似するデータラベルから絵描き歌のベースとなる指示を作成.

課題については, 文章が英語であることなどが挙げられていた.

絵描き歌の生成という機能を 2 段階に分けると, スケッチから歌詞への変換と, 歌詞から曲への変換に分けられる. したがって事前調査として, 「画像から文章への変換」や「作曲の自動化」などについて調べた.

2.1 画像から文章への変換

先行研究では, スケッチから推測した単語列をもとに絵描き指示を生成していた.

他の方法では, スケッチではないが画像からキャプションを生成する研究 ([1] など) がある. この方針の場合, 英語ではなく日本語で絵描き歌用のデータセットを作成する必要があるため, 先行研究の方法よりも難しいと思われる.

2.2 作曲の自動化

機械学習による作曲も, 入力と出力によって問題の種類が様々あった. メロディー以外にも音楽の要素は多数存在し, 生成方法も小節単位や時間単位, 音符単位などがあり, データフォーマットも 音声信号, MIDI, ピアノロール, テキスト形式があった.

音楽に関する知識がないので, 「作曲の自動化」についてはかなり厳しいと感じた.

3 方針

「作曲の自動化」はかなり難しいと分かったため, 「画像から文章への変換」に絞る方針にしたい. 先行研究で改良できそうな点として,

- 文章をストローク単位からオブジェクト単位に変える
- 描き順を絵描き歌用に最適化する.

の 2 つを思いついた.

1 点目は, 絵描き歌の単位をストロークとしていたが, 実際の絵描き歌はオブジェクト単位となっていたことから, スケッチを木構造のようなグループに分割できるとより自然な絵描き歌になると考えた. 手法として, ストロークの組み合わせの中から最適なものを, スケッチデータベースから検索することで, スケッチをオブジェクトの集合で表せられる.

2 点目, 先行研究は描きはじめの位置や順番などを忠実に守っていたが, 実際の例を確認すると絵描き歌にする際には強引な書き順になることも多かった. 改良点として, よく似ているオブジェクト順に書き順を並べ替えることでさらに自然になると考えた. この類似度指標には先行研究の SSII や描き順を考慮しない指標を用いたい.

4 今後の予定

こんな感じの方針でいいのか相談したいです. とりあえずは藤井さんの追実験のしようと思っています.

参考文献

- [1] Oriol Vinyals, Alexander Toshev, Samy Bengio, and Dumitru Erhan. Show and tell: A neural image caption generator. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 3156–3164, 2015.