

合成音声品質の評価基準に関する研究

Study on evaluation criteria of synthesized speech quality

ジョニダ¹⁾

指導教員 亀田弘之²⁾, 渡邊紀文³⁾, 喜多義弘²⁾, 相田紗織²⁾

- 1) 東京工科大学・コンピュータサイエンス学部 バイオ・情報メディア研究科 思考と言語研究室
2) 東京工科大学 コンピュータサイエンス学部
3) 武蔵野大学 工学部

キーワード：合成音声，評価基準，環境感，親切感

1. はじめに

言語は人間相互の意思疎通のための道具であるため、これをコンピュータで処理することが出来ればその社会的意義は大きい。このような観点から、音声処理・自然言語処理の名の下、多くの研究がなされてきている。より質の高いサービスを実現するためには、なおも合成音声の品質を高める必要がある。それは必ずしも多くはない。しかも、「品質」を評価する標準は、人によって違うので、多彩な声を作成できる音声合成技術は、無限の可能性を持っており、非常に高い研究価値がある。このような考えに立脚し、本研究では、「個々のユーザの主観的観点から、合成音声の良し悪しを判断するための評価基準を提案する」ことである。

2. 関連研究

2.1 評価方法に関する研究[1]

NTT 通信トラヒック品質プロジェクトによれば、音声品質評価法には大きく分けて「主観的評価」と「客観的評価」の二つに分類することができる。

主観品質を評価する基本的な方法は主観品質評価技術である。主観品質評価技術の最大の特徴は、お客様の感じる品質を視聴覚心理実験によって直接測定するという点である。そのため、主観評価は

サービスの主観品質を評価するうえで最も信頼できる方法である。

主観品質評価技術に対して、主観評価により得られる値と同等の値を、音声や映像の物理的特徴から推定する方法を客観品質評価技術と呼ぶ。

2.2 評価項目に関する研究[2]

CTTS (Corpus-based Text-To-Speech) の技術動向によれば、音声品質は評価項目に大きく依存する。

「個々のユーザの主観的観点から、合成音声の良し悪しを判断するための評価基準を提案する」ことを目指し、関連研究の一部評価基準を参考し、新たな評価基準を作り、提案する。

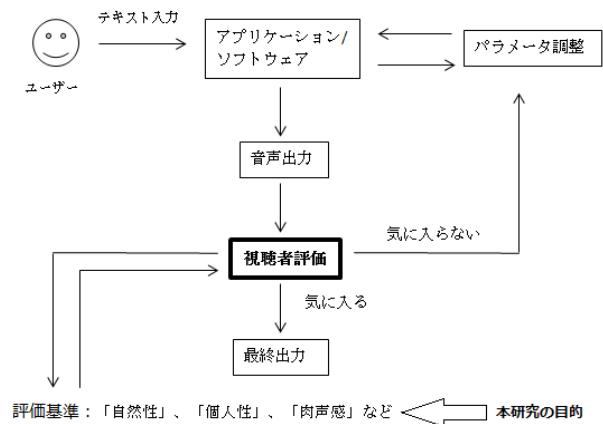


図1. システム概念図

「環境感」

環境感は元々音質領域で使うものではなく、映画やアニメなどの方がよく使う用語である。本来は画面中の人と全般の割合、調和性などを評価する基準だが、本研究では、音声と画面の同調度を評価する基準とする。つまり、聴覚と視覚の一致を保証するように、特定な雰囲気に調整するものととられる。

[2] 渡辺聰・岩木健・兼安勉・三木敬 『ヨーパスベース音声合成とその応用』 沖テクニカルレビュー 73(2), 62-65, 2006-04

環境感の評価基準（5点満点）

1	2	3	4	5
場所は全く想像できない	大体の場所が分かる	想像できるが、雰囲気があまり感じられない	ある程度の雰囲気が感じられる	まるで本当にその場にいる

表1. 環境感の採点基準

「親切感」

親切感は音声の総合評価を評価する基準である。つまり、音楽を聞いてる時、その音声を通して、作者の見せたいものが想像が想像できるように音声を調整するものととられる。

親切感の評価基準（5点満点）

1	2	3	4	5
まるで舞台の裏側で見てる	ある程度の目を合わせることができる	作者が見せたいものが分かる	作者の気持ちが分かる	まるで本人と対話するように感じる

表2. 親切感の採点基準

5. まとめ

目的

本研究の目的は、音声に関する主観評価項目を提案することである。

提案項目

- 1) 「環境感」
- 2) 「親切感」

今後の予定

更に提案項目及び採点基準を詳しく検討する。

参考文献

- [1] <http://www.ntt.co.jp/qos/technology/sound/02.html> NTT 通信トラヒック品質プロジェクト