

どれくらい正しく聞き取れるか分かります ～正解文が不要な音声認識率推定技術～

どんな研究

音声認識システムの性能を示す音声認識率は、認識結果文と正解文を比較して算出します。しかし事前に人が音声を聞いて正解文を作成する必要があるため、そのコストが非常に高いという問題点がありました。本展示では、**正解文を用いずに音声認識率を推定する技術**を紹介します。

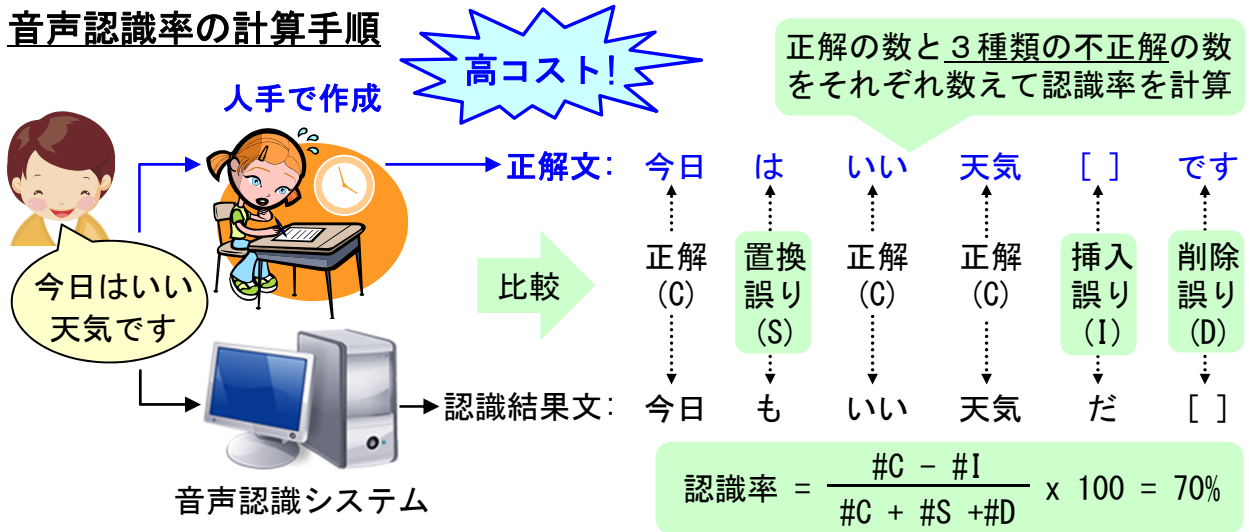
どこが凄い

認識率は認識結果文中の各単語を正解または**3種類の不正解**のいずれかに分類することで計算します。本研究では、正解文を用いずに、この分類を確率的に行う**誤りタイプ分類技術**を開発しました。本技術によれば、**認識率が高精度に推定可能**であることを確認しました。

目指す未来

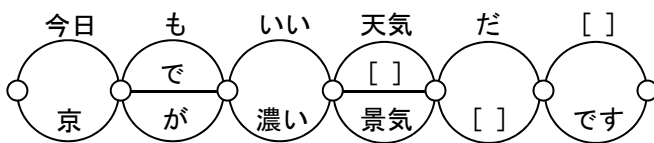
音声認識技術は着実に進展し応用例も増えていますが、話者や環境などの変動への頑健性は十分ではありません。本技術を用いれば、音声認識技術の導入判断やシステムの自己診断と性能改善が**低コスト**で可能になるなど、**音声認識技術の応用先の拡大に貢献**すると期待できます。

音声認識率の計算手順

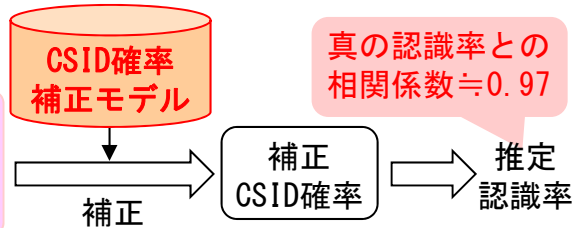
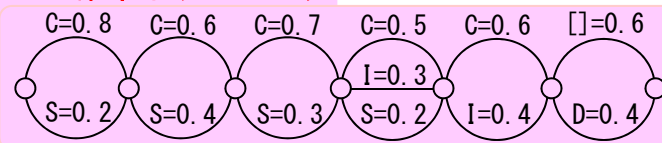


誤りタイプ分類 (CSID確率推定) に基づく認識率推定

認識結果文 (含, 対立候補)



CSID確率ネットワーク



正解文が不要で低コストかつ高い認識率推定精度 (言語依存性なし)

関連文献

- [1] A. Ogawa, T. Hori, A. Nakamura, "Recognition rate estimation based on word alignment network and discriminative error type classification," in *Proc. IEEE Workshop on Spoken Language Technology (SLT)*, pp. 113-118, 2012.
- [2] A. Ogawa, T. Hori, A. Nakamura, "Discriminative recognition rate estimation for n-best list and its application to n-best rescoring," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pp. 6832-6836, 2013.

連絡先

小川 厚徳 (Atsunori Ogawa) メディア情報研究部 信号処理研究グループ
E-mail : ogawa.atsunori[at]lab.ntt.co.jp ([at] の部分を @ に置き換えてください)