

感情認識合成のための大規模感情音声コーパス「JTES」

感情認識によりAIや音声認識で高精度、表現豊かな対話を実現

概要

AIなどを利用した音声対話システムが普及している。その音声対話システムの多くは、発話内容の言語的情報のみを扱い対話が行われている。そのため対話内容によっては、対話が成り立たない場合や対話の不自然さを感じる場合といった課題があった。

JTES (Japanese Twitter-based Emotional Speech)は入力された音声の声色等から感情を推定する「感情認識」、感情音声の入力に対して音声の認識を行う「感情音声認識」、感情の乗った音声を合成する「感情音声合成」を実現する音声対話システムに利用するための汎用感情音声コーパスである。これにより、感情認識によりAIや音声認識で高精度、表現豊かな対話を実現する。

実施例条件

話者：一般話者（演技経験を問わない）

感情：喜び・怒り・悲しみ・平常の各50文を男女各50名、計100名が発話した20,000発話（23.5時間）

技術データシート

関連特許（整理番号）：- (S20-017、東北大学整理番号:S20180004)

著作者： 能勢 隆、伊藤 彰則、千葉 祐弥

実施例

感情音声によるモデル学習の感情音声認識における有効性の検証することにより、感情音声認識による本音声コーパスの評価を行った。

指標	精度・効果	備考
感情認識率	73.5%	ディープラーニングにより人間による認識率(75%)に近い認識精度を達成
感情音声認識単語誤り率(WER)	8.0%	高い精度が得られ、評価と学習の感情が一致するとき精度が上がる傾向
感情音声合成	有効	客観・主観評価から感情音声を用いて合成することの有効性有り

連絡先

株式会社 東北テクノアーチ

TEL 022-222-3049

FAX 022-222-3419

問い合わせは[こちら](#)からお願いします。