

山川 仁子
Kimiko YAMAKAWA

松井 知子
Tomoko MATSUI

菊池 英明
Hideaki KIKUCHI

板橋 秀一
Shuichi ITAHASHI

どんな研究？

音声に関する研究において、音声データは欠かせないものとなっており、国内外で様々な種類のコーパスが構築されています。しかし、大量の音声コーパスの中から、利用目的に合ったコーパスを選択するのは容易ではありません。本研究では、コーパスの特徴を表す属性項目を提案し、複数のコーパスの中から利用者が利用目的にあったものを簡単に直感的に選ぶことができるようなコーパス検索の方法を提案します。

なにがわかる？

コーパス特徴を表す属性項目の提案と、特徴属性を用いた多次元尺度構成法による音声コーパス特徴の可視化を提案します。これにより、それぞれのコーパスがもつ特徴を明示しながら、複数音声コーパス間の類似点や差異を直感的・視覚的に表現することが可能になります。また、提案手法を用いたコーパス検索システムの構築を進めています。利用目的に合ったコーパスを容易に選択することが可能になるだけでなく、分野別・国別のコーパス特徴分析など、コーパスに関する研究の促進にもつながります。

1. コーパス検索の背景

音声コーパス

- 音声研究には音声コーパスが必要不可欠
- 多種多様なコーパスの構築・公開

コーパスの多様化は...

膨大なコーパスの中から、利用目的に合ったものを選ばなければならない

利用者

選択肢の拡大
利用選択の範囲が広がる

現在のコーパス検索方法

- テキストによる取り扱いコーパリストからの選択
- 機関によって異なるコーパス情報

利用目的にあったコーパスを探し、複数コーパスを比較することは困難

研究目的 簡便なコーパス検索方法の検討

1. コーパス特徴を表す属性項目の提案
2. 複数のコーパスの関連性を可視化

2. コーパスの特徴を表す属性項目

表1. コーパス属性項目(9属性53項目)

#	属性名	項目数	内容
1	入力デバイス	7	卓上マイク, 接話マイク, ピンマイク, 固定電話, 携帯電話, 放送, 不明
2	入力環境	5	オフィス・教室, 防音室, 騒音下, 車内, 不明
3	発話スタイル	4	連続音声, 孤立音声, 非母語音声, 不明
4	発話モード	5	対話・会話, 会議, 講演, 読み上げ, 不明
5	データモード	9	標本周波数(≤8kHz, ≤16kHz, >16kHz), 筋電図, 音声パラメータ, 画像, パラトグラム, MRI画像, 不明
6	言語	4	単一言語, 多言語, 方言, 不明
7	目的	14	音声分析, 音声合成, 音声認識, 話者認識, 言語認識, 数字認識, マルチモーダル, ロバスト, 高齢者, 子供, 非母語話者, 不特定話者, 非音声(騒音など), 不明
8	話者数	3	男性話者数, 女性話者数, 全体話者数
9	音響特徴量	2	SN比(平均), SN比(分散)

属性項目のベクトル変換

属性(1)~(7): 当てはまるか否かを{1,0}で表現
話者数属性(8): 延べ話者数を正規化した量的データで表現
音響特徴量(9): 全音声ファイルの10%のSN比を, 正規化した量的データで表現

音声ファイルごとの対数フレームパワーの分布を音声と雑音による2混合のガウス分布として推定する手法[1]を使用

3. 複数コーパス間の類似度可視化

解析手続き

1. ベクトル変換したコーパス特徴属性からユークリッド距離による類似度行列を計算。
2. Torgerson 法による計量的MDSを適用してコーパスの空間配置を導出
3. k-means法によるクラスタ分析

表2. 解析対象コーパス (NII-SRCより配布の26種類)

#	corpus name	#	corpus name
1	連続音声DB (PASL-DSR)	14	CENSREC-4
2	筑波大多言語音声DB	15	メディア教育利用 (ERJ)
3	方言音声 (GSR-JD)	16	メディア教育利用 (JRF)
4	RWCP 対話音声 (SP96)	17	千葉大MapTask
5	RWCP 対話音声 (SP97)	18	宇都宮大パラ言語DB
6	RWCP ニュース音声 (SP99)	19	電総研単語DB(ETL-WD)
7	RWCP 会議音声 (SP01)	20	鶴岡方言音声DB 91
8	音声対話DB (PASD)	21	新聞記事読上げ音声DB(JNAS)
9	CIAIR 子供の声 (CIAIR-VCV)	22	新聞記事読上げ高齢者音声(S-JNAS)
10	CENSREC-1	23	NTT・東北大 親密度別音声DS(FW03)
11	CENSREC-1-C	24	NTT・東北大 親密度別音声DS(FW07)
12	CENSREC-2	25	NTT幼児音声DB (Infant)
13	CENSREC-3	26	日本音響学会 連続音声DB(JIPDEC)

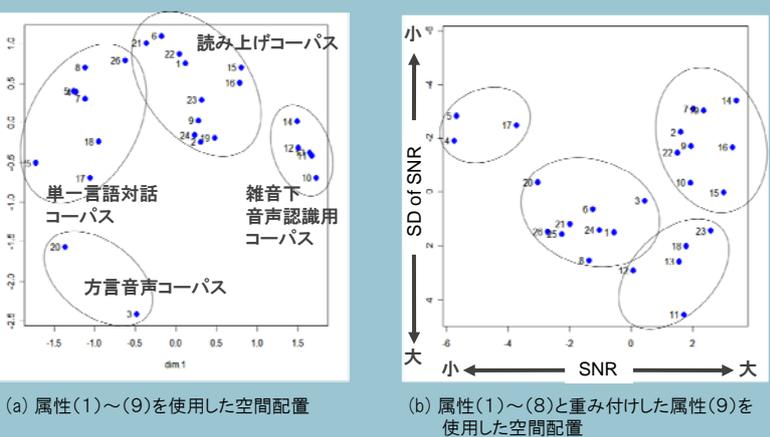


図1 コーパス属性を用いたNII-SRC 26コーパスの空間配置

4. コーパス検索システム

提案手法を用いた音声コーパス検索システムの構築

- Web上でのコーパス検索
- コーパス関連情報の提示
- 他機関のコーパス情報も包括

■ 近日, 公開検証実験を実施

5. まとめ

コーパス特徴属性による複数コーパス間の関係性を可視化

◎ 多角的にコーパス特徴を表現
コーパス利用者のニーズや利用目的, コーパス作成者の作成意図を反映

◎ 複数コーパスの比較が容易
特徴属性を反映した布置により, 直感的な比較が可能

【引用文献】
[1]T. H. Dat, K. Takeda, F. Itakura, "On-line Gaussian mixture modeling in the log-power domain for signal-to-noise ratio estimation and speech enhancement," Speech Communication, 48(11),1515-1527, 2006.

【連絡先】
国立情報学研究所
音声メディアグループ
<http://research.nii.ac.jp/src/src@nii.ac.jp>