

会議音声データベースを利用した 愛想笑い検出器の開発



沼尾研究室
Numao Laboratory



竹之内翔太郎 沼尾雅之（電気通信大学）

はじめに

- 笑い方によって、心情的意味合いは異なってくる
→ 含み笑いのような「愛想笑い」の存在に注目



愛想笑い

- 愛想笑いの検出
→ 会議における状況把握に役立てるのでは？

愛想笑いをリアルタイムで検出するための
アプリケーションを開発したい



課題点

1. 話声やノイズなどと区別して、笑い声を検出すること
2. 検出された笑い声が、愛想笑いかどうかを判定すること

★ 音声から笑い声の分類まで行った既存研究は数少なく、その手法もバッチ処理を前提としている[1]

→ 上記の課題点を考慮して、愛想笑い検出手法を提案

[1] H.Tanaka, N.Campbell. "Classification of social laughter in natural conversational speech", Computer Speech & Language(28), 2014.

会議音声データベースの特徴

- コーパス付属の書き起し文に注目
 1. 笑い声などに対する擬音語での書き起し
→ 擬音語表現のパターンから、笑いの種類を定義
 2. 会議内で生じたノイズについても書き起こしデータがある

会議中の笑い声に対する分析

- 辞典[2]を参考にして、笑い声に対する擬音語表現のパターンを次の3つに分類した

- A) 普通の笑い: 「アハハ」「ハッハハ」 など
B) 含み笑い: 「ウフフ」「フッフッフ」 など
C) 照れ隠しの笑い: 「エヘヘ」「ヘッヘヘ」 など

→ A～Cのパターンに一致した笑い声に対してラベルを付けた

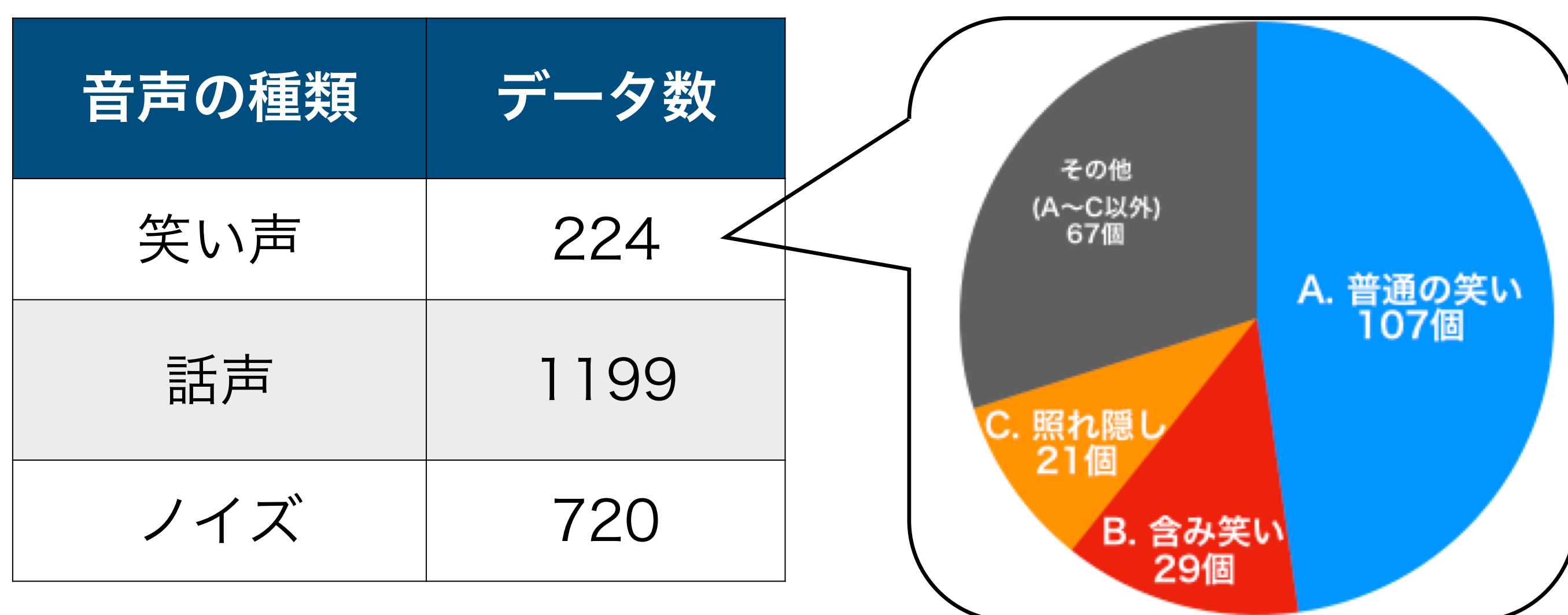


図1 会議音声データベース内の音声に対する分析

[2] 「笑う・笑い」の擬音語・擬態語: <http://www.jlogos.com/d010/4381608.html> (2018/11/27 閲覧)

提案手法

- リアルタイム認識を想定した、愛想笑いの検出手法を提案
・ 音声認識エンジン[3]を使い、マイクからの認識を実装

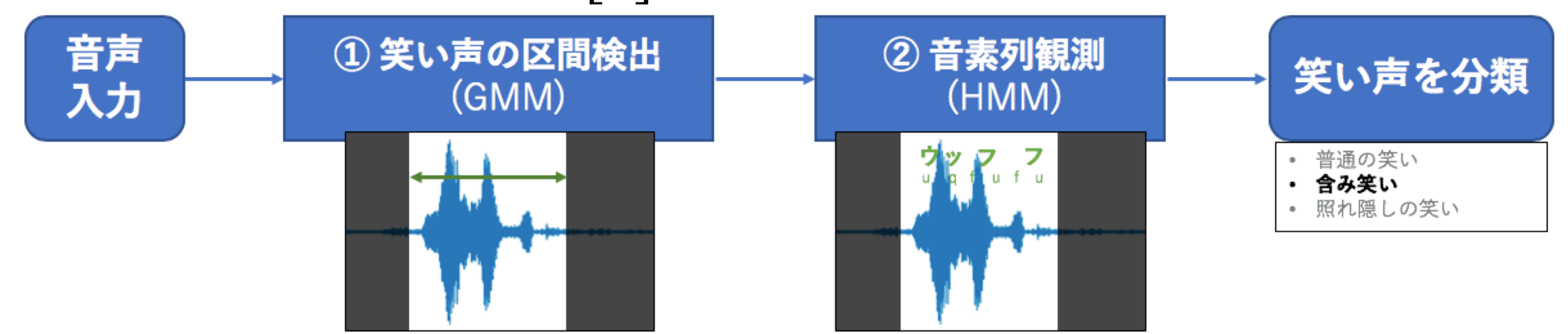


図2 愛想笑いの検出までの過程

① GMMによる笑い声の区間検出について

- GMMによる尤度を用いて、笑い声の区間を検出
- 話声やノイズの棄却も、ここで行っている

② HMMによる音素列観測について

- 笑い声のような非言語音は、学習時のラベリングが難しい
- そこで、笑い声に対する擬音語での書き起しデータを活用
→ 笑い声の音素定義・音素ラベルの位地決りを自動化

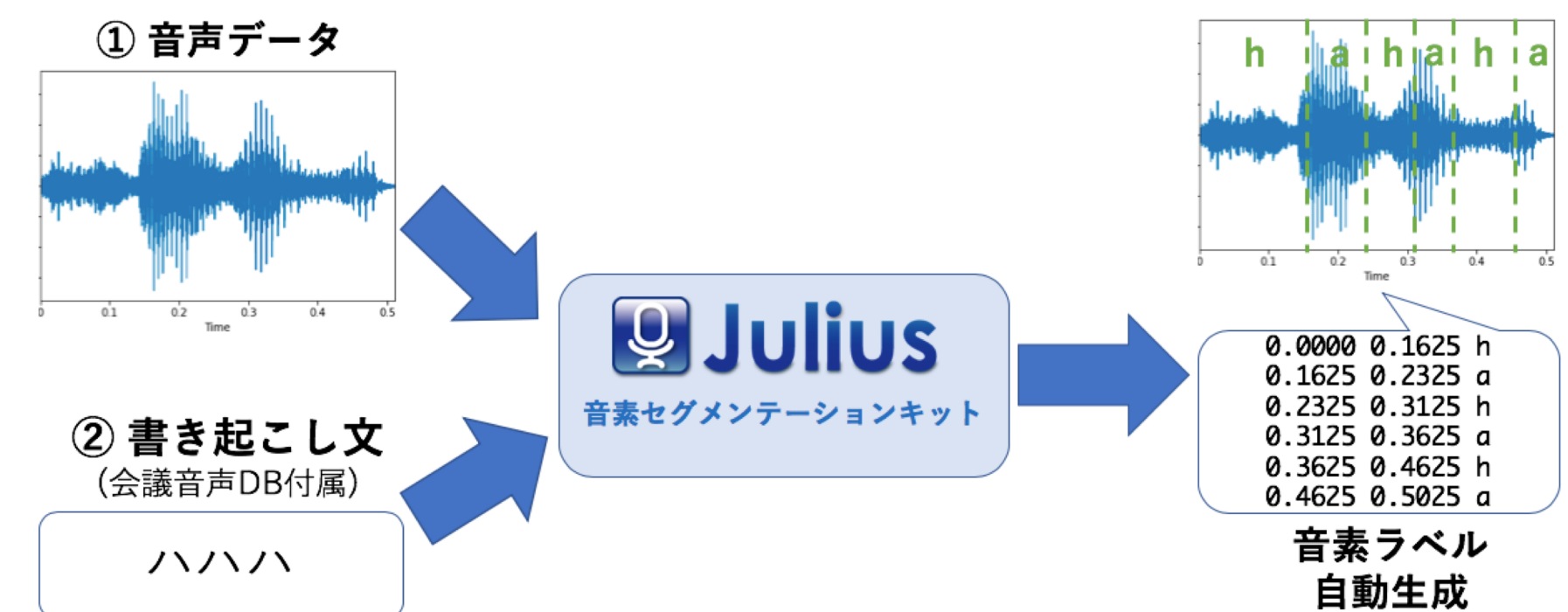


図3 笑い声に対する自動ラベリング

[3] A.Lee, et al. "Julius --- An Open Source Real-Time Large Vocabulary Recognition Engine", EUROPEECH2001: the 7th European Conference on Speech Communication and Technology, 2001.

評価

① GMMに対する評価

- 笑い声の検出と、笑い声以外の棄却精度に関して評価する
 - ・ 図2における、①の部分に相当
 - ・ 笑い声の音声に対し、正しく1回分の笑いを検出できたら正解とする（複数回に分かれたら不正解とした）
- 5-fold Cross-Validationでの検証を行った
 - ・ 笑い声の検出率は、75.8%となった
 - ・ GMMによる、話声とノイズの棄却精度は良かった

音声の種類	誤検出率
話声	0.42 %
ノイズ	5.69 %

② HMMに対する評価

- 観測した音素列から、含み笑いなどの検出精度を評価する
- 図1における、A～Cに属する笑い声同士で分類を行った
 - ・ 10-fold Cross-Validationでの検証
 - ・ 結果は下表のようになり、BとCの分類精度が良くなかった

		Predicted		
		A	B	C
True	A. 普通の笑い	74	17	16
	B. 含み笑い	10	13	6
	C. 照れ隠しの笑い	7	3	11

③ リアルタイム性について

- 音声入力の終了を検知した後、認識処理を実行している
 - ・ 具体的には、図2における①と②の処理を行う
- 認識処理に要する時間は、1～3秒
→ 音声に来てから、高々3秒ほどで認識ができる

おわりに

- 笑い声の認識に適した、認識時のパラメータに対するチューニング方法について検討を行いたい
- 笑い声の学習データ量が少なかったため、データを増やして精度が向上するかどうかについても検証したい
- 笑い声に対する擬音語での書き起し → 音素定義に役立ったため、HMMによる非言語音認識に活用できるという期待を得た
 - ・ 生活行動認識への応用を目的に、他の非言語音(咳やいびきなど)の認識にも拡張できるかどうか、今後検討したい

謝辞: 本研究では、国立情報学研究所より提供している会議音声データベース(RWCP-SP01)を利用した。