

側音化構音音声認識についての検討

研究概要

従来の音声認識システムは正確に発音した音声を対象としており、構音障害を持つ人のような発音に癖のある音声については対象としていないため認識精度が落ちてしまう。

本研究では構音障害の症状の一部である「側音化構音」について注目し、その症状に適した音声認識システムの構築を検討する。

側音化構音

舌が口蓋に接しているため音声・呼気が口腔の中央から出ず側面から出てしまう症状。摩擦音や破裂音などが側面音に変化してしまう。例として下表の赤字は「●」の音へと変化しやすい。

構音位置	両唇音		歯音・歯茎音		硬口蓋音		軟口蓋音		声門音	
	無声	有声	無声	有声	無声	有声	無声	有声	無声	有声
破裂音	p	b	t	d			k	g		
破擦音			ts	dz						
			tʃ	dʒ						
摩擦音	ɸ		s	z	ç			ɣ	h	
			ʃ	ʒ						
通鼻音		m		n		ɲ		ŋ		
弾音				r						
側面音			●							

話者適応

特定の話者の音声認識において認識精度を向上させる方法として、特定の話者の音声データのみで学習を行い、その話者専用の音響モデルをつくる手法がある。しかし作成するためには特定の話者の音声データを大量に入手する必要があり困難である。

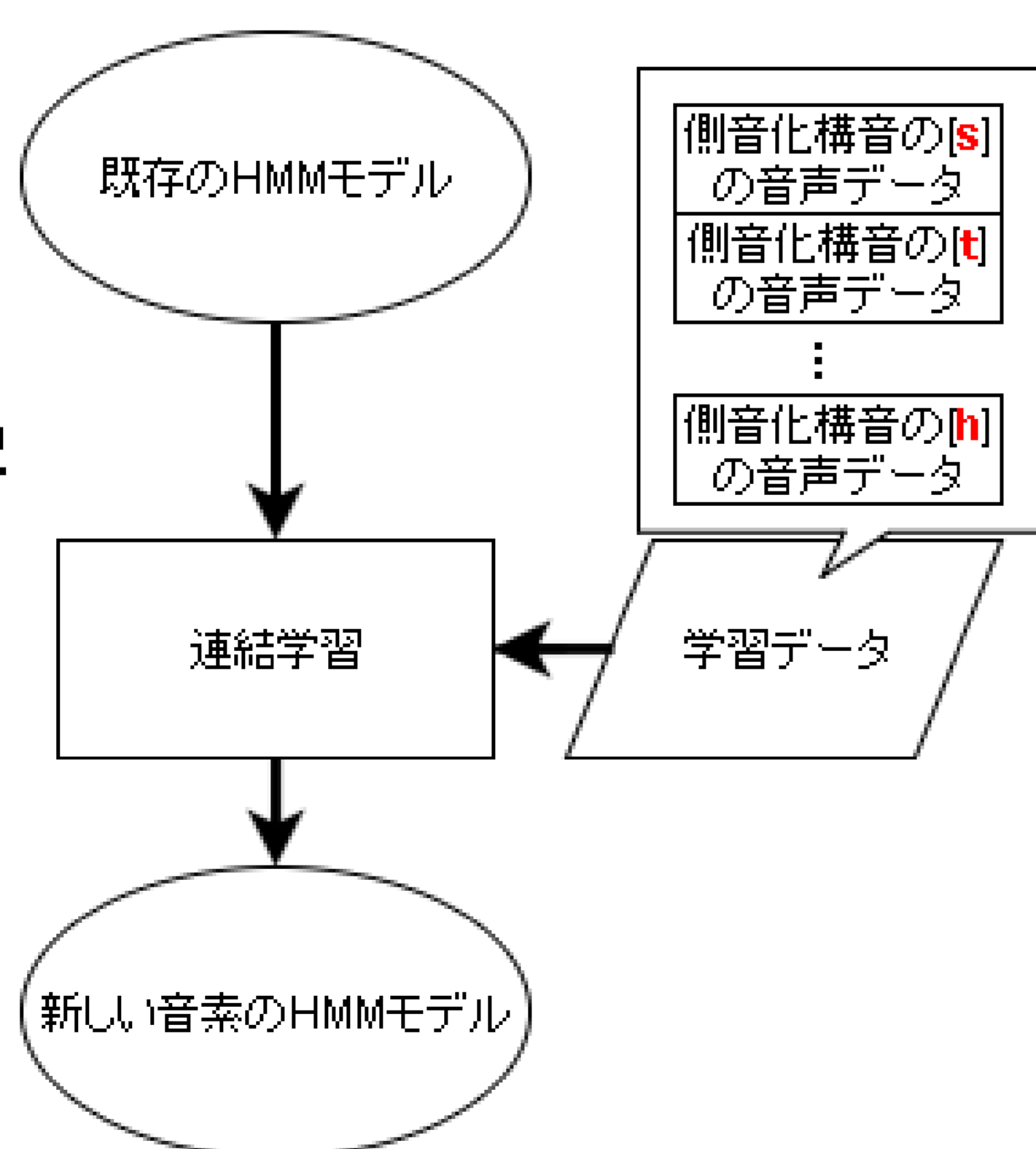
そこで不特定話者の学習データで作成された音響モデルを利用し、特定の話者の音声データで連結学習を行うことで音響モデルを作成する手法が話者適応である。この手法では比較的少量の音声データで特定話者の音響モデルを作成でき、認識精度の向上が望める。

提案手法

側音化構音では特定の子音が別の子音に変化することに注目し、新しい音素を追加した音響モデルとそれに伴う発音辞書の作成を行う。側音化構音で変化の子音と「新しい音素」の対応関係に基づき、下の処理を行う。

音響モデル

新しい音素に対応した音素について、全てを学習データとし適応学習を行い、「新しい音素」のHMMモデルを作成する



発音辞書

エントリ	出力	音素列
すいか+名詞	[すいか]	suika
つい+動詞	[つい]	tsui
ひらめ+名詞	[ひらめ]	hirame
⋮		
すいか+名詞	[すいか]	sth u i k a
つい+動詞	[つい]	ts u i
つい+動詞	[つい]	sth u i
ひらめ+名詞	[ひらめ]	h i r a m e
ひらめ+名詞	[ひらめ]	sth i r a m e

側音化構音で変化する音素を「新しい音素」に置き換えた発音の追加する

実験・結果

実験者が503文を読み上げ、それについて適応学習を行わない場合、各音素に話者適応を行った場合、今回提案した手法の場合でそれぞれの手法で認識結果を比較した。

今回のデータについて、提案した手法は通常の話者適応より僅かながら認識精度が向上した。

手法	認識率(%)
適応学習なし	33.24
通常の話者適応	42.20
今回提案した手法	43.21