Internal Use Only (非公開)

TR-I-0100

ベクトル量子化話者適応化の研究

A Study of VQ-based Speaker Adaptation

中村　哲
Satoshi NAKAMURA

1989. 8

概要

ベクトル量子化を用いた話者適応化の研究を行った。セパレートベクトル量子化、ファジィベクトル量子化、ファジィベクトル量子化のファジィ級関数を確率とみなしてコードベクトルの対応づけを推定するファジィヒストグラム推定アルゴリズム、ファジィベクトル量子化で表された未知話者の入力音声を連続的に標準話者の空間に写像するファジィ写像を提案した。この方法を、HMM音韻認識、時間遅れ神経回路網(TDNN)による音韻認識、声質変換に適用し、有効性を明らかにした。また、HMMにおいて、話者間の調音様式を適応化する話者重畳型HMMの提案を行い音韻認識における有効性を示した。さらに、ニューラルネットによる非線形写像アルゴリズム、線形写像アルゴリズムとのスペクトル写像精度の比較を行い、ベクトル量子化話者適応アルゴリズムの有効性を示した。

ATR 自動翻訳電話研究所
ATR Interpreting Telephony Research Laboratories

© ㈱ATR 自動翻訳電話研究所　1989
© 1989 by ATR Interpreting Telephony Research Laboratories

目次

1. まえがき 1
2. セパレートベクトル量子化に基づく話者適応化 1
3. ファジィベクトル量子化に基づく話者適応化 1
4. ベクトル量子化話者適応のHMM音韻認識への適用 1
5. ベクトル量子化話者適応の時間遅れ神経回路網による音韻認識への適用 2
6. 話者重畳型HMMを用いた調音様式の話者適応化 2
7. 教師つき話者適応化における写像方法の比較 3
8. 発表リスト 4

1. まえがき

ベクトル量子化を用いた話者適応化の研究を行った。ベクトル量子化は、話者の音声の空間の離散表現を可能とするため、話者間ではこの離散点の対応づけを求めることで話者適応化が実現できる。本稿では、セパレートベクトル量子化、ファジィベクトル量子化による高精度化手法を、HMM、時間遅れ神経回路網(TDNN)に適用して有効性を示した。さらに話者間での音韻の調音様式の違いに起因する時系列パターンの違いに対処する話者重畳型HMM、ニューラルネットによる非線形写像、線形写像との写像精度の比較について述べた。

2. セパレートベクトル量子化に基づく話者適応化

ベクトル量子化の手法が、スペクトル空間の離散表現になるという観点から、未知話者のスペクトログラムを標準話者のスペクトログラムに正規化する話者適応方法について検討を行った。ここで述べる正規化方法では、ベクトル量子化により有限個のベクトルで話者毎のスペクトル空間を表現した後、予め学習により求めた話者間でのベクトルの対応関係を用いて正規化を行う。ここでは、従来のスペクトルとパワーによる複合スペクトル歪み尺度に基づいた単一のコードブックによるベクトル量子化に対し、スペクトルとパワー毎にコードブックを用いるセパレートベクトル量子化を導入することで、話者内のベクトル量子化歪み及び話者間のスペクトログラム正規化誤差を小さくすることを試みた。評価をスペクトル歪みおよび認識実験で行い、セパレートベクトル量子化が有効であることを明らかにした。

3. ファジィベクトル量子化に基づく話者適応化

ベクトル量子化が、スペクトル空間の離散表現になるということを利用して、未知話者のスペクトログラムを標準話者のスペクトログラムに写像する話者正規化方式について検討を行った。ベクトル量子化を用いたスペクトログラムの話者間での正規化にファジィベクトル量子化の考えを導入する。ファジィベクトル量子化は、既存の各コードベクトルに対する帰属度を用いて入力ベクトルを表現するため、量子化歪を小さくできる。さらにファジィベクトル量子化にk近傍則を導入し高速化、高精度化を行い、スペクトログラムの話者間正規化に適用した。また、話者間の対応関係を示す変換コードブックの学習において、ファジィベクトル量子化の級関数を確率として用い、変換コードブックの高精度化を行う。音韻バランス216単語、男女計3名分の音声データに対し、スペクトル歪、及び話者間正規化の正規化誤差を用いて評価を行い、ファジィベクトル量子化を用いたスペクトログラム正規化方式の有効性を明らかにした。

4. ベクトル量子化話者適応のHMM音韻認識への適用

ベクトル量子化話者適応化アルゴリズムをHMM音韻認識に適用する。HMMに適用する際には、動的特徴を考慮したセパレートベクトル量子化、ファジィベ

クトル量子化、ファジィヒストグラム、ファジィマッピングを用いる。さらに、HMMとの効率的な整合のために、対応づけヒストグラムを標準話者のファジィ級関数としてファジィHMMを計算する話者適応アルゴリズムを用いる。音韻バランス216単語、重要語5240単語を用いて有声破裂音/b,d,g/および全音韻の音韻認識実験を男女計3名の話者について行なった結果、次の事柄が確かめられた。(1)動的特徴を考慮したセパレートベクトル量子化を用いることにより有声破裂音の認識率が6.4%改善できる、(2)ファジィベクトル量子化を用いることにより有声破裂音の認識率が3.4%改善できる、(3)ファジィヒストグラムを用いることにより話者適応化の学習に必要な単語数を100単語から25単語に削減しても認識率の低下を0.4%に抑えられる。また、有声破裂音の認識率は、男性間で83.1%、男女間で76.5%で、従来法[M.Feng et al. ICASSP 88]との比較では11.7%の認識率の改善となること、全音韻の認識では、男性間で75.6%、男女間で71.8%で、上位3位までの累積認識率では、男性間、男女間いずれの場合にも約91%を達成できることがわかった。

5. ベクトル量子化話者適応の時間遅れ神経回路網による音韻認識への適用

ベクトル量子化話者適応アルゴリズムを時間遅れ神経回路網(TDNN)を用いた音韻認識へ適用する。TDNNへの適用にあたってはTDNNの入力パラメータの比較、ベクトル量子化を用いた場合のTDNNの構成の検討を行う。入力パラメータとしては、FFT、LPC分析によるスペクトルとケプストラム、自己相関係数の比較を行う。また、ベクトル量子化を用いた場合のTDNNの構成としてパラメータ入力、ベクトル量子化の符号を入力とするTDNNを検討し話者適応化を適用する。音韻バランス216単語、重要語5240単語、国際会議申し込みに関する会話文を用いて有声破裂音/b,d,g/の認識実験を男女計4名の話者について行った。この結果、(1)TDNNの入力パラメータの形式については、周波数領域のパラメータが優れている。周波数領域で表現されていれば、FFTでも、LPCの分析方法でも大差がない。(2)ベクトル量子化をTDNNに適用する場合、符号列入力形式のTDNNよりもパラメータを入力とするTDNNが優れている。ベクトル量子化としては、ファジィベクトル量子化を用いることでベクトル量子化による認識率の劣化を改善できる。(3)ベクトル量子化話者適応アルゴリズムをTDNNに適用した結果、男性間と男女間の平均で78.7%の認識率が得られ、話者適応化が有効に行えることが明かとなった。

6. 話者重畳型HMMを用いた調音様式の話者適応化

HMMにおける話者適応化を改善するために複数話者HMM学習を用いた話者適応化方法を提案する。話者間における発声の違いとしては、(1)静的なスペクトルの違い(2)調音様式に起因する時系列パターンの違いがある。現在までに我々は、ベクトル量子化話者適応化法を用いて、(1)静的なスペクトルの違いについての適応化を行い有効性を示した。ここでは、(2)の時系列パターンの違いを適応化

するために、複数話者HMM学習を用いた話者適応化法を提案する。この方法は、基準標準話者に加えてさらに複数の標準話者を用意し、基準話者のスペクトルでありながら複数の話者から得られた種々の時系列パターンを1つのモデルの中に含むHMM(話者重畳型HMMと呼ぶ)を生成する。有声破裂音の認識による評価を行った結果、話者重畳HMMにより82.5%の認識率が得られ本方法の有効性を確認した。

7. 教師つき話者適応化における写像方法の比較

これまでに話者適応をフレーム毎のスペクトルの写像としてとらえベクトル量子化話者適応により話者適応化ができることを示した。一方、近年、学習アルゴリズムの確立によりニューラルネットワークの研究が進み音声の分野にも適用され始めた。話者適応化においても文献(5,6)にニューラルネット適用の試みがみられる。本稿では、教師つき話者適応化の枠組みの中で、ベクトル量子化話者適応アルゴリズムにおけるファジィ写像、線形写像、ニューラルネットを用いた非線形写像の写像精度の比較評価を行う。この結果、線形写像では近傍数を3として区分線形部分空間の写像を行なうときベクトル量子化話者適応化アルゴリズムより1.6%低い写像精度となる、ニューラルネットよる話者適応化ではいずれの場合も5.5%から6%程度ベクトル量子化話者適応化アルゴリズムより低い写像精度となることが示された。

8. 発表リスト

本研究に関する大会講演、論文の発表リストを次に示す。

1.大会講演

著者(発表者)	題名	発表先	年月	ページ
中村　哲 鹿野清宏	ベクトル量子化を用いた スペクトログラムの正規化	電子情報通信学会 音声研究会	1987.6	9
中村　哲 鹿野清宏	ベクトル量子化を用いた スペクトログラムの正規化	音響学会全国大会	1987.10	18
阿部匡伸 中村　哲 鹿野清宏 桑原尚夫	ベクトル量子化による声質変換	音響学会全国大会	1987.10	20
中村　哲 鹿野清宏	ベクトル量子化を用いた スペクトログラムの正規化	電子情報通信学会 部門別全国大会 シンポジウム	1987.11	22
中村　哲 鹿野清宏	ファジィベクトル量子化を用いた スペクトログラムの正規化の検討	電子情報通信学会 音声研究会	1988.2	24
阿部匡伸 中村　哲 鹿野清宏 桑原尚夫	ベクトル量子化による声質変換	電子情報通信学会 音声研究会	1988.2	31
中村　哲 鹿野清宏	ファジィベクトル量子化を用いた スペクトログラムの正規化	音響学会全国大会	1988.3	37
阿部匡伸 中村　哲 鹿野清宏 桑原尚夫	ベクトル量子化による声質変換の評価	音響学会全国大会	1988.3	39
中村　哲 花沢利行 鹿野清宏	ファジィベクトル量子化に基づく話者適応化 のHMM音素認識による評価	音響学会全国大会	1988.10	41
中村　哲 花沢利行 鹿野清宏	ベクトル量子化話者適応アルゴリズムのHMM 音韻認識による評価	電子情報通信学会 音声研究会	1988.12	43
中村　哲 花沢利行 鹿野清宏	ベクトル量子化話者適応化の HMM音韻認識による評価	音響学会全国大会	1989.3	52
阿部匡伸 中村　哲 鹿野清宏 桑原尚夫	入力スペクトル情報の利用による声質変換法 の高度化	音響学会全国大会	1989.3	54
中村　哲 鹿野清宏	時間遅れ神経回路網(TDNN)における入力パラ メータの評価と話者適応化	電子情報通信学会 音声研究会	1989.6	56
中村　哲 鹿野清宏	教師つき話者適応化における写像方法の比較	音響学会全国大会	1989.10	65

1.大会講演

著者(発表者)	題名	発表先	年月	ページ
中村　哲 服部浩明 鹿野清宏	複数話者HMM学習を用いた話者適応化の音韻認識による評価	音響学会全国大会	1989.10	67
花沢利行 中村　哲 川端　豪 鹿野清宏	ベクトル量子化話者適応アルゴリズムのHMM文節認識による評価	音響学会全国大会	1989.10	69

2.学術論文

著者	題名	番号	年月	ページ
阿部匡伸 中村　哲 鹿野清宏 桑原尚夫	Voice Conversion through Vector Quantization	ICAASP 88	1988.4	71
中村　哲 鹿野清宏	Spectrogram Normalization Using Separate Vector Quantization	7th FASE Symposium	1988.8	75
中村　哲 鹿野清宏	セパレートベクトル量子化を用いた スペクトログラムの正規化	音響学会論文誌	1988.8	83
中村　哲 鹿野清宏	ファジィベクトル量子化を用いた スペクトログラムの正規化	音響学会論文誌	1989.2	92
中村　哲 鹿野清宏	Speaker Adaptation Applied to HMM and Neural Networks	ICASSP 89	1989.5	101
中村　哲 花沢利行 鹿野清宏	ベクトル量子化話者適応のHMM音韻認識 への適用	音響学会論文誌	査読終了	
阿部匡伸 中村　哲 鹿野清宏 桑原尚夫	Voice Conversion through Vector Quantization	音響学会論文誌 投稿中	1989.4	
中村　哲 鹿野清宏	Conparison of Input Parameters for Time Delay Neural Network and Implementation of Speaker Adaptation	ICASSP 90 投稿中	1989.8	
中村　哲	A Comparative Study of Spectral Mapping for Speaker Adaptation	ICASSP 90 投稿中	1989.8	
服部浩明 中村　哲	Supplementation of Articulatory Speaker Variation for Speaker Adaptation on HMM	ICASSP 90 投稿中	1989.8	
中村　哲 鹿野清宏	時間遅れ神経回路網(TDNN)における入力パラ メータの評価と話者適応化	電子情報通信学会 論文誌 準備中	1989.8	
中村　哲 鹿野清宏	教師つき話者適応化における写像方法 の比較評価	電子情報通信学会 論文誌 準備中	1989.8	
中村　哲 服部浩明 鹿野清宏	話者重畳型HMMによる調音変動 の話者適応化	電子情報通信学会 論文誌 準備中	1989.8	

3.特許

発明人	発明の名称	出願年月日	出願番号	備考
中村　哲 鹿野清宏	スペクトログラムの正規化方式	62. 6.24	62-156958	
中村　哲 鹿野清宏	ベクトル量子化を用いた音声認識方式	62. 6.24	62-156959	
中村　哲 鹿野清宏	ベクトル量子化を用いた音声認識方式	62. 6.24	62-156960	
阿部匡伸 中村　哲 鹿野清宏 桑原尚夫	音質変換方式	62.10.9	62-255498	
中村　哲 鹿野清宏	スペクトログラムの正規化方式	63.3.17	62-156958	