

TR-I-0210

ニューラルネットワークによる音声認識の研究

--- 発表論文集 ---

A Study of Speech Recognition

Using Neural Networks

--- A Collection of Papers ---

沢井 秀文

Hidefumi SAWAI

1991. 3.29.

概要

本テクニカル・レポートは、著者がATRにおいて発表した論文(共著を含む)の一覧である。テクニカル・レポートTR-I-0209: "Connectionist Large Vocabulary Continuous Speech Recognition" に最近の成果について全般的な解説を述べたので、合わせて御利用下さい。

ATR 自動翻訳電話研究所

ATR Interpreting Telephony Research Laboratories

© (株)ATR 自動翻訳電話研究所 1991

© 1991 by ATR Interpreting Telephony Research Laboratories

目次

1. まえがき	1
2. 発表論文リスト	2
2.1. 大会講演	2
2.1.1. 日本音響学会	2
2.1.2. 電子情報通信学会等	3
2.2. 電子情報通信学会音声研究会	4
2.3. 学術論文	5
2.4. 国際会議発表	6
2.5. 解説記事その他の発表	8
3. テクニカル・レポート一覧	8
4. 出願特許一覧	10
5. 発表論文	11

1. まえがき

本テクニカルレポートは、ニューラルネットワーク(NN)を用いた音声認識処理と、言語処理を統合した大語彙連続音声認識システムの最近の成果について述べた姉妹編のテクニカルレポートTR-I-0219 “Connectionist Large Vocabulary Continuous Speech Recognition”に関連する論文を含め、著者が発表した論文の一覧である。

大語彙連続音声認識システムの音声認識部は、日本語の24音素(即ち18子音 /b/, /d/, /g/, /p/, /t/, /k/, /m/, /n/, /N/, /s/, /sh/, /h/, /z/, /ch/, /ts/, /r/, /w/, /y/ と5母音 /a/, /i/, /u/, /e/, /o/ と促音 /Q/ または無音)を特別なセグメンテーション技術を必要とせずに、単に入力音声の中をスキャンするだけで自動的にスポッティングすることができる時間遅れ神経回路網(TDNN)から成っている。

一方、言語処理部は予測LRパーザから成る。このLRパーザは文脈自由文法から自動的に生成されたLRパージングテーブルに従って動作し、バックトラック無しに左から右へ処理が行なわれる。予測音素と一連のTDNN音素出力結果との時間整合は、DTWマッチング法により行われる。この統合された認識システムをTDNN-LR法と呼ぶ。TDNN-LR法を用いた大語彙と連続音声認識が優れた特定話者認識性能を示す。特に発声速度の適応化には少量の学習トークンによる追加学習の効果が顕著であることが判明した。

さらにTDNN-LR認識システムの拡張として、幾つかの新しい成果についても述べている。即ち、新しいNNアーキテクチャを提案し、発声変動に対し従来のTDNNに比べロバストな音素認識が実現できること、標準話者と未知話者との間のスペクトル・セグメントの静的・動的な写像をNNによるマッピング関数として実現する話者適応化技術について、また従来マルチ・スピーカー用に提案されていたNNアーキテクチャを改良し、不特定話者認識に適した新しいアーキテクチャを提案し、特定話者並みの音素認識率が得られた等である。これらの成果は、今後、TDNN-LR認識システムに統合されて、よりロバストな不特定話者向きの音声認識システムへと発展していくことが期待される。

2. 発表論文リスト

2.1. 大会講演

2.1.1. 日本音響学会

著者(発表者)	題 名	発 表 先	発表年月
沢井 秀文 A. Waibel 鹿野 清宏	時間遅れ神経回路網による音韻スポットティングの検討	日本音響学会秋季研究発表会 2-P-11, pp223-224.	昭和63年10月 1988.10.
A. Waibel 沢井 秀文 鹿野 清宏	Phoneme Recognition by Modular Construction of Time-Delay Neural Networks	日本音響学会秋季研究発表会 2-P-12, pp225-226.	昭和63年10月 1988.10.
P. Haffner 沢井 秀文 A. Waibel 鹿野 清宏	Fast Back-Propagation Learning Methods for Large Phonemic Neural Networks	日本音響学会春研究発表会 1-6-14, pp27-28.	平成元年3月 1989.3.
宮武 正典 沢井 秀文 鹿野 清宏	全音韻を統合した時間遅れ神経回路網(TDNN)による音韻スポットティング	日本音響学会春研究発表会 2-P-24, pp277-278.	平成元年3月 1989.3.
沢井 秀文 宮武 正典 鹿野 清宏	時間遅れ神経回路網による階層的な音韻スポットティング	日本音響学会春研究発表会 2-P-25, pp279-280.	平成元年3月 1989.3.
南 泰浩 宮武 正典 沢井 秀文 鹿野 清宏	TDNN音韻スポットティングと拡張LRパーザを用いた文節音声認識	日本音響学会秋季研究発表会 3-1-11, pp97-98.	平成元年10月 1989.10.
J.C.Dang 沢井 秀文	音韻認識のための決定論的なボルトマンマシン	日本音響学会秋季研究発表会 1-1-23, pp45-46.	平成元年10月 1989.10.
宮武 正典 沢井 秀文 鹿野 清宏	時間遅れ回路網による音韻スポットティングの改良	日本音響学会秋季研究発表会 1-1-25, pp49-50.	平成元年10月 1989.10.
中村 雅己 田村 震一 宮武 正典 沢井 秀文	ニューラルネット開発用ワークベンチシステム-ネットワークエディタおよびモニタ機能について	日本音響学会秋季研究発表会 2-P-26, pp181-182.	平成元年10月 1989.10.
福沢 圭二 沢井 秀文 杉山 雅英	ニューラルネットワークによる恒等写像を用いた話者適応	日本音響学会秋季研究発表会 1-8-16, pp31-32.	平成2年9月 1990.9.

著者(発表者)	題 名	発 表 先	発表年月
杉山 雅英 福沢 圭二 沢井 秀文 嵯峨山 茂樹	ニューラルネットによる集合間写像の教師なし学習	日本音響学会秋季研究発表会 2-P-10, pp149-150.	平成2年9月 1990.9.
沢井 秀文	TDNN-LR文節音声認識システムにおける追加学習の効果	日本音響学会秋季研究発表会 2-P-11, pp151-152.	平成2年9月 1990.9.
沢井 秀文	時間・周波数変動に強い時間遅れ神経回路網(TDNN)	日本音響学会秋季研究発表会 2-P-12, pp153-154.	平成2年9月 1990.9.
南 泰浩 沢井 秀文 中川 正雄	TDNNの構造の違いによる特性の比較	日本音響学会春季研究発表会 1-5-5, pp13-14.	平成3年3月 1991.3.
沢井 秀文 中村 悟 福沢 圭二 杉山 雅英	ニューラルネットワークによる不特定話者音声認識へのアプローチについて	日本音響学会春季研究発表会 1-5-17, pp37-38.	平成3年3月 1991.3.

2.1.2. 電子情報通信学会等

著者(発表者)	題 名	発 表 先	発表年月
P Haffner 沢井 秀文 A. Waibel 鹿野 清宏	Fast Back-Propagation Learning Methods for Neural Networks in Speech Recognition	電子情報通信学会春季全国大会, SA-1-1, pp1-329-330.	平成元年3月 1989.3.
宮武 正典 沢井 秀文 鹿野 清宏	時間遅れニューラルネットワークを用いた音韻スポッティング法	電子情報通信学会89年春期大会, SA-1-4, pp1-335-336.	平成元年3月 1989.3.
沢井 秀文	時間遅れ神経回路網(TDNN)による音声認識	平成2年電気関係学会関西支部連合大会 S14-2	平成3年10月 1991.10.

2.2. 電子情報通信学会音声研究会

著者(発表者)	題 名	発 表 先	発表年月
沢井 秀文 A. Waibel 宮武 正典 鹿野 清宏	モジュール構成ニューラルネット ワークのスケールアップによる音 韻認識	電子情報通信学会音声研究会 SP88-105, pp73-80.	昭和63年12月 1988.12.
宮武 正典 沢井 秀文 鹿野 清宏	連続音声中の音韻スポッティング のためのTDNN構成法	電子情報通信学会音声研究会 SP89-32, pp63-68	平成元年6月 1989.6.
J.C.Dang 田村 震一 沢井 秀文	シフト不変型決定論的ボルツマン マシンによる音声認識	電子情報通信学会音声研究会 SP89-98, pp1-6.	平成元年10月 1989.10.
南 泰浩 沢井 秀文 宮武 正典 鹿野 清宏	TDNN音韻スポッティングと予測 LRパーザを用いた大語彙単語音声 認識	電子情報通信学会音声研究会 SP89-99, pp7-13.	平成2年6月 1990.6.
中村 悟 沢井 秀文	不特定話者音素認識のための ニューラルネットアーキテクチャ の検討	電子情報通信学会音声研究会 SP90-61, pp33-40.	平成2年11月 1990.6.

2.3. 学術論文

著者(発表者)	題 名	発 表 先	発表年月
沢井 秀文 宮武 正典 A. Waibel 鹿野 清宏	連続音声認識のための時間遅れ神経回路網を用いた音韻/音節スポットティング	電子情報通信学会D-II分冊「新しい音声処理」特集号、招待論文、vol. J72-D-II, No.8, pp1151-1158.	平成元年8月 1989.8.
A. Waibel 沢井 秀文 鹿野 清宏	Modularity and Scaling in Large Phonemic Neural Networks	IEEE Transaction on ASSP, vol.37, No. 12, pp - .	平成元年12月 1989.12.
沢井 秀文 宮武 正典 A. Waibel 鹿野 清宏	Spotting Phonemes and Syllables for Continuous Speech Recognition Using Time-Delay Neural Networks	Systems and Computers in Japan, vol. 21, No.9, pp71-79.	平成2年9月 1990.9.
宮武 正典 沢井 秀文 鹿野 清宏	時間遅れ神経回路網(TDNN)による音韻スポットティングのための学習法とその効果	電子情報通信学会D-II分冊, vol. J73-D-II, No.5, pp699-706.	平成2年5月 1990.5.
南 泰浩 沢井 秀文 宮武 正典	時間遅れ神経回路網(TDNN)による音韻スポットティング法と予測LRパーザを用いた大語彙単語音声認識	電子情報通信学会D-II分冊, vol. J73-D-II, No.6, pp788-795.	平成2年6月 1990.6.
沢井 秀文 南 泰浩 宮武 正典 A. Waibel 鹿野 清宏	Connectionist Approaches to Large Vocabulary Continuous Speech Recognition	電子情報通信学会英文誌「連続音声認識・理解小特集号」(掲載予定)	平成3年7月 1991.7.

2.4. 国際会議での発表

著者(発表者)	題 名	発 表 先	発表年月
沢井 秀文 A. Waibel 宮武 正典 鹿野 清宏	Spotting Japanese CV-Syllables and Phonemes Using TimeDelay Newral Networks	ICASSP (International Conference on Acoustics, Speech and Signal Processing) '89, Glasgow, England.	平成元年5月 1989.5.
A. Waibel 沢井 秀文 鹿野 清宏	Consonant Recognition by Modular Construction of Large Phonemic Time-Delay Neural Networks	ICASSP (International Conference on Acoustics, Speech and Signal Processing) '89, Glasgow, England.	平成元年5月 1989.5.
沢井 秀文 A. Waibel P Haffner 宮武 正典 鹿野 清宏	Parallelism, Hierarchy, Scaling in Time-Delay Neural Networks for Spotting Japanese Phonemes/CV Syllables	IJCNN (International Joint Conference on Neural Networks) '89, IEEE/INNS主催, Washington D.C., U.S.A.	平成元年6月 1989.6.
P Haffner A. Waibel 沢井 秀文 鹿野 清宏	Fast Back-Propagation Learning of Large Phonemic Database	European Conference on Speech Communication and Technology, Paris, France.	平成元年9月 1989.9.
宮武 正典 沢井 秀文 南 泰浩 鹿野 清宏	Integrated Training for Spotting Japanese Phonemes Using Large Phonemic Time-Delay Neural Networks	ICASSP (International Conference on Acoustics, Speech, and Signal Processing) '90, Albuquerque, U.S.A.	平成2年4月 1990.4.
沢井 秀文	The TDNN-LR Large-vocabulary and Continuous Speech Recognition System	1990 International Conference on Spoken Language Processing, Kobe	平成2年11月 1990.11.
沢井 秀文	TDNN-LR Continuous Speech Recognition System Using Adaptive Incremental TDNN Training	ICASSP(International Conference on Acoustics, Speech and Signal Processing) '90, Toronto, Canada	平成3年5月 1991.5. (発表予定)
沢井 秀文	Frequency-time-shift-invariant time-delay Nural Networks for Robust Continuous Speech Recognition	ICASSP(International Conference on Acoustics, Speech and Signal Processing) '90, Toronto, Canada	平成3年5月 1991.5. (発表予定)
杉山 雅英 沢井 秀文 Alex Waibel	Review of TDNN (Time-Delay Neural Network) Architectures for Speech Recognition	ISCAS (International Symposium on Circuits and Systems) '91, Singapore	平成3年6月 1991.6. (発表予定)
福沢 圭二 沢井 秀文 杉山 雅英	Segment-based Speaker Adaptation by Neural Network	IEEE, SP Society, Workshop on Neural Networks for Signal Processing, Princeton, U.S.A.	平成3年9月 1991.9. (発表予定)

著者(発表者)	題 名	発 表 先	発表年月
沢井 秀文 中村 悟	Time-Delay Neural Network Architectures for High-performance Speaker-independent Recognition	2nd European Conference on Speech Communication and Technology, Genova, Italy	平成3年9月 1991.9. (発表予定)

2.5.解説記事その他の発表

著者(発表者)	題 名	発 表 先	発表年月
鹿野 清宏 川端 豪 畑崎 香一郎 沢井 秀文 A. Waibel	Speech Recognition Research at ATR	文部省重点研究第2回国際シンポジウム	昭和63年11月 1988.11.
沢井 秀文	TDNN の 展 開 及 び ICASSP,89/IJCNN,89会議報告	文部省科研費総合研究研究討論会	平成元年7月 1989.7.
田村 震一 沢井 秀文 中村 雅己 鹿野 清宏	ニューラルネットワークを用いた音声処理	テレビジョン学会誌, vol.43, No.9, pp935-943, 1989.	平成元年8月 1989.8.

3.テクニカル・レポート

著 者	題 名	番 号	年 月
A. Waibel 沢井 秀文 鹿野 清宏	Modularity and Scaling in Large Phonemic Neural Networks	TR-I-0034	昭和63年8月 1988.8.
P.Haffner A. Waibel 沢井 秀文 鹿野 清宏	Fast Back-Propagation Learning Methods for Neural Networks in Speech	TR-I-0058	昭和63年11月 1988.11.
鹿野 清宏 片桐 滋 田村 震一 中村 雅己 A. Waibel 沢井 秀文 P.Haffner 東倉 洋一 E.McDermott 船橋 賢一	ニューラルネットワークの音声情報処理への応用	TR-I-0063	昭和63年12月 1988.12.
南 泰浩 宮武 正典 沢井 秀文 鹿野 清宏	TDNN音韻スポッティングと拡張LRパーザを用いた文節音声認識	TR-I-0085	平成元年7月 1989.7.

著 者	題 名	番 号	年 月
沢井 秀文 宮武 正典 A. Waibel 鹿野 清宏	時間遅れ神経回路網を用いた音韻/音節スポットティング	TR-I-0089	平成元年7月 1989.7.
沢井 秀文 宮武 正典 A. Waibel P.Haffner 鹿野 清宏	Parallelism,Hierarchy,Scaling in Time-Delay Neural Networks for Spotting Phonemes and CV-Syllables	TR-I-0090	平成元年7月 1989.7.
宮武 正典 沢井 秀文 鹿野 清宏	時間遅れ神経回路網(TDNN)による音韻スポットティングのための効果的学習法	TR-I-0103	平成元年8月 1989.8.
中村 雅巳 田村 震一 宮武 正典 沢井 秀文	ニューラルネット開発用ワークベンチシステムネットワークエディタおよびモニタ機能について	TR-I-0113	平成元年9月 1989.9.
Jean-Claude Dang 田村 震一 沢井 秀文	Shift-invariant Deterministic Boltzmann Machines for Phoneme Recognition	TR-I-0130	平成元年12月 1989.12.
南 泰浩 宮武 正典 沢井 秀文 鹿野 清宏	TDNN音韻スポットティングと拡張LRパーザを用いた大語彙単語音声認識	TR-I-0144	平成2年2月 1990.2.
南 泰浩 沢井 秀文	TDNNの構造の音韻認識率、シフトインバリエント性への影響	TR-I-0145	平成2年2月 1990.2.
中村 悟 沢井 秀文	時間遅れ神経回路網を用いた不特定話者の音韻認識	TR-I-0178	平成2年9月 1990.9.
南 泰浩 沢井 秀文	発話変動にロバストなTDNNの検討	TR-I-0183	平成2年10月 1990.10.

4. 出願特許

発明者	題 名	出願年月日	出願番号	備考
沢井 秀文 宮武 正典 鹿野 清宏	ニューラルネットによる連続音声 認識方式	平成元年10月3日	1-259359	
沢井 秀文	ニューラルネットワークを用いた 音声認識方式	平成2年10月1日		
福沢 圭二 沢井 秀文 杉山 雅英	ニューラルネットワークによる話 者適応化方式	平成2年10月 日		
沢井 秀文	ニューラルネットによる不特定話 者音声認識方式	平成3年3月 日		