

科学研究費助成事業 研究成果報告書

平成 26 年 6 月 20 日現在

機関番号：32689

研究種目：基盤研究(B)

研究期間：2010～2013

課題番号：22300060

研究課題名(和文)大規模音声ドキュメントを対象とする超高速検索エンジンの構築

研究課題名(英文)Designing an ultra-high speed search engine for big data of spoken documents

研究代表者

新田 恒雄(Nitta, Tsuneo)

早稲田大学・グリーンコンピューティングシステム研究機構・教授

研究者番号：70314101

交付決定額(研究期間全体)：(直接経費) 10,600,000円、(間接経費) 3,180,000円

研究成果の概要(和文)：ビッグデータ中の音声ドキュメントから任意のキーワードを、実時間で検索する技術を開発した。

研究実施にあたっては、(1)未知語を含む音声を高精度に音素列へ変換する技術、(2)曖昧性を含む音素列からキーワードを高速に検索する技術の二つに焦点をあてた。

(1)では、双対空間で音素特徴を効率よく抽出すると共に、多層パーセプトロンで調音素性を抽出し、音素を高精度に識別する方式を開発した。(2)では、接尾辞配列に基づき反復深化探索を行う方式をベースに、調音素性間の距離計算を用いた連続DP、およびキーワード分割アルゴリズムを実装することで、検索精度、検索速度、記憶容量の三つの課題を同時に克服できることを示した。

研究成果の概要(英文)：Fast spoken term detection from big data has been developed. In the development, we have focused on (1) accurate speech-to-phoneme conversion even though there is an out-of-vocabulary word in an utterance, and (2) fast spoken term detection even though there is ambiguity caused by speech recognition errors.

In (1), after extracting features of phonemes in a dual space, a multi-layer perceptron extracts articulatory features, then a subsequent phoneme classifier discriminates phonemes with high accuracy. In (2), we have implemented an iterative deepening search based on suffix arrays, a continuous DP matching, and a keyword division algorithm. As a result, we could solve three issues of search accuracy, search speed, and size of index at the same time.

研究分野：総合領域

科研費の分科・細目：情報学・知覚情報処理、知能ロボティクス

キーワード：音声情報処理 ビッグデータ 高速検索 サフィックス・アレー 音素認識 調音素性抽出 ディープニューラルネット 双対空間

1. 研究開始当初の背景

研究開始時点では、インターネット上の大規模データベース検索がテキストを対象に始まっていた。続いて、検索対象を音声ドキュメントへと拡張することへの期待が高まりつつあった。しかし、当時の認識技術は、大語彙連続音声認識 (Large Vocabulary Continuous Speech Recognition; LVCSR) を用いて音声をテキストに変換するため、検索対象に未知語を含むと変換誤りを生じ、検索精度が低下する問題があった。誤りを多く含む音声認識結果に対して、音素(列)などの部分単語(subword)を並行して検出する試みも見られたが、この場合、音素識別精度が課題とされていた。一方、当時の標準的音声ドキュメント検索方式は、2,000 時間の音声に対してキーワード検索に 3 秒程度掛かっており、大規模コーパスに対する音声検索は不可能であった。

2. 研究の目的

本研究は、大規模音声ドキュメントから任意キーワードを、実時間で検索する技術を開発することを目指してスタートした。この実現に向け、

- (1) 任意単語を含む音声を高精度に音素列へ変換する技術、および
- (2) 曖昧性を含む音素列からキーワードを高速に検索する技術

の二つの技術開発を課題とした。(2)では、未知語を含む大規模音声ドキュメントに対して、1 秒内にキーワードを検索できることを目標とした。

3. 研究の方法

(1) 高速化へのアプローチ: テキスト検索分野で研究されていた接尾辞配列に着目し、曖昧性を持つ音声へ適用できるように改良を図る。音声から音素列への変換は、誤り(置換・付加・脱落)を含むため、誤りと時系列の伸縮を許容する動的時間伸縮 (Dynamic Time Warping; DTW, 我が国では別名 Dynamic Programing; DP) を利用する必要がある。一方、DP の導入は接尾辞配列の木構造探索範囲を一挙に増大させる。そこで、検索システムは結果をスコア順に上位から順に表示できれば十分なことに着目し、探索の閾値を逐次制御する「反復深化探索」を導入することとした。

(2) 高精度化へのアプローチ: 調音素性 (Articulatory Feature; AF), すなわち調音運動を陽に表す特徴を音声から高精度に抽出する方式を開発し、高い音素認識性能を得る方式へ特化することとした。

4. 研究成果

平成 22 年度から平成 25 年度に至る研究成果を、検索アルゴリズムの改良に関連する、(1)検索精度、(2)高速化、(3)記憶容量低減の三点、および音素認識アルゴリズムの改良に関連する、(4) 調音素性抽出 DNN (Deep Neural Network), (5) 双対空間 (Dual space) を利用した DNN の負担軽減と性能改良の二点を中心に説明する。

(1) 検索精度の改良

接尾辞配列は、音声の全ての音素に対するインデックスを格納した配列を、音素列の接尾辞順にソートしたものである。ソート済みのデータ構造を持つため、検索キーワードを効率的に見つけることができる。図 1 に音素列(図では character で示す) /abracadabra/ に対して接尾辞配列を構築した例を示す。

Characters	a	b	r	a	c	a	d	a	b	r	a
インデックス	1	2	3	4	5	6	7	8	9	10	11

↓ Extract suffixes

インデックス	Suffix	インデックス	Suffix
1	abracadabra	11	a
2	bracadabra	8	abra
3	racadabra	1	abracadabra
4	acadabra	4	acadabra
5	cadabra	6	adabra
6	adabra	9	bra
7	dabra	2	bracadabra
8	abra	5	cadabra
9	bra	7	dabra
10	ra	10	ra
11	a	3	racadabra

Sort →

Array={1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11} Suffix Array={11, 8, 1, 4, 6, 9, 2, 5, 7, 10, 3}

図 1 接尾辞配列の構築

図中のインデックスはその接尾辞(suffix)が入力音素列の何番目から始まるかを表す。そこで、キーワード/bra/を検索する場合、インデックスから元の音素列を参照しながら二分探索すると 9 と 2 の位置に/bra/が出現することが分かる。ソートされたインデックスのみを保持すれば良いため、必要なデータ領域を小さくすることができる。また、インデックスから元の音素列を参照すると、任意長の音素列の出現位置を検索できるという特性を持つ。

接尾辞配列では完全一致検索を想定している。このため、誤認識を含む音声認識結果を対象とするには曖昧検索技術を導入する必要がある。我々は山下らが提案した端点固定 DP マッチングに変えて、同じキーワードを複数検索可能な連続 DP を採用して実装した。検索途中で、累積距離が閾値を越えると探索は打ち切られる。

DP マッチングの実行中は、閾値内のすべてのパスが保持されるため、閾値が大きいと探索空間および処理時間が指数関数的に増加する。閾値は検索キーワードの長さ按比例して増加させる必要があるため、キーワードが

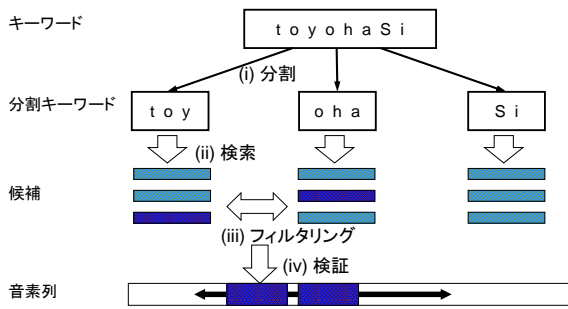


図2 キーワード検索の流れ

長いと処理時間が急激に増加する。そこでこの問題を解決するため、キーワードを分割し、分割されたキーワードを元のキーワードの代わりに検索する手法を導入した(図2参照)。

具体的には、分割した場合と分割しない場合で同一の検索結果が得られるよう、(i)分割、(ii)検索、(iii)フィルタリング、(iv)検証の各ステップを設け、この中で閾値を一定の数式下で制御すればよいことを見出し実装した。

図3にキーワード検索の主要方式間における性能比較を示す。実験は Intel Core i7-2600 プロセッサ 3.4GHz, メインメモリ 8GB を搭載した PC で行った。評価音声ドキュメントは、NTCIR-9 フォーマルランの CSJ コーパス CORE 講演 39 時間と ALL 講演 606 時間である。基本的性能を確認するため、(A) 高速検索の基礎技術となっている N-gram インデックスベースの検索法、(B) NTCIR-9 が提供するベースライン、(C) 累積距離 T をキーワード長 l で正規化した値を閾値 t とする提案法の簡易版、および (D) 短いキーワードの出現を抑制するため式 $score = 1 / [t / (l)^{1/2}]$ を共通の閾値として実装した提案手法の4つを比較した。再現率は正解箇所のうち検索できた箇所の割合、適合率は検索結果の中の正解の割合である。なお後述する最大 F 値は、再現率と適合率の調和平均である。

図3の再現率-適合率曲線から分かるように、提案手法は音素 3-gram インデックスを用いた手法を上回っている。調音素性の利用は検索精度向上に効果があった。提案法の (C) と (D) を比較すると、(D) の score を用いた手法は (C) の t を用いた手法を上回っており、score によるキーワード長補正が有効なことを確認できた。

(2) 高速性の比較評価

提案手法は、CORE 講演、ALL 講演を対象とした時、1.56msec で検索結果を出力した (F 値最大の条件で測定)。データ数の少ない 39 時間の CORE 講演と、606 時間の ALL 講演で検索速度が同じになった理由は、F 値が最大となる score の値が ALL で高かったため

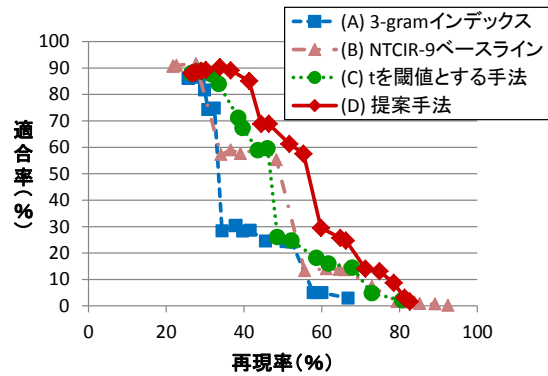


図3 検索基本性能の比較

ある。NTCIR-9 参加グループの中で、本手法は参加グループの中でも最速グループに入っていた。ベースラインの連続 DP マッチングを用いる手法は ALL 講演の検索に 548 秒を要しており、このことから接尾辞配列の導入が大幅な高速化を可能にしたことが分かる。

(3) 記憶容量の比較評価

NTCIR-9 参加グループの中で、高速な検索性能を示したグループは、メモリあるいは二次記憶を多く消費していた。39 時間の CORE 講演を例にとると、600MB~3.5GB (さらにリスコアリングで精度を出す場合は、25.6GB の二次記憶領域が必要なケースも) と、大きな容量を必要とする。

提案手法は、メモリや二次記憶を消費しないという特性がある。このため、長時間の ALL 講演でも 241MB、さらに精度向上のために 2 種の音声認識エンジンを利用した場合で、メモリ使用量 469MB と低く抑えることができた。

(4) 調音素性(AF)抽出 DNN

音声信号は発音器官の動作により生成されるため、音声学分野の知見を通して、古くから生成過程が明確になっていた。音素認識では、音声を調音素性系列へ精度よく変換できると、分類器への負担が軽減され、同時に分類性能が向上すると期待される。図4に、

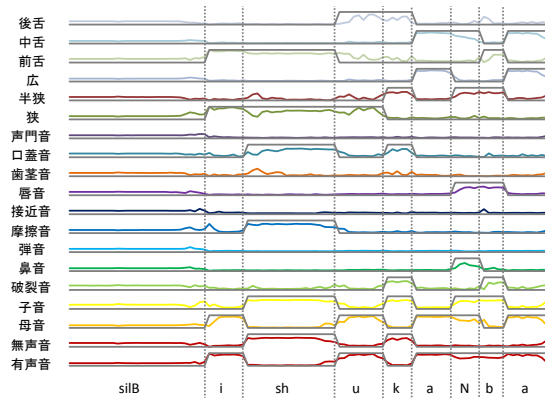


図4 調音素性の抽出例(「一週間ばかり」)

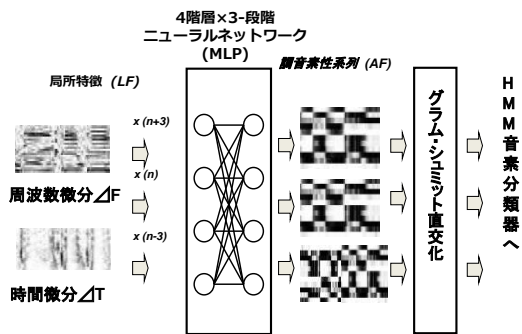


図 5 調音素性 (AF) 抽出の過程

図 5 の調音素性 (AF) 抽出器を用いて、音声を調音素性系列へ変換した例を示す。図 5 は、AF 抽出器をニューラルネットワーク (多層パーセプトロン (Multi-layer Perceptron); MLP) で構成している。全体は 4 層から成る MLP を 3 段、即ち 12 層から成っており、近年、DNN (Deep Neural Network) と呼ばれる技術の先駆けと言える。入力は、24 チャンネル帯域通過フィルタの出力 (パワースペクトル) を時間と周波数方向に微分した局所特徴 (local feature; LF) を採用している。LF は、 3×3 の空間演算子で、畳み込み演算を実現したもので、図に示すように ΔT はパワースペクトル系列の過渡部 (主に子音部) を、また ΔF は定常部 (主に母音部) を表現している。これも近年、畳み込みニューラルネットワーク (Convolutional NN) として盛んに使用される方式の先駆けと言える。図の最初の MLP-1 は教師信号に、調音素性 (例えば音素 /b/ であれば、有声、破裂、口唇が 0.9、それ以外の素性には 0.1) を与えて学習している。二段目の MLP-2 は調音素性の時間的文脈を学習し、更に最終段 MLP-3 は調音素性の時間方向微分 (速度 Δ) と二次微分 (加速度 $\Delta\Delta$) を入力に与えて学習している。先に述べた図 4 は、以上の処理を通して得た調音素性系列を示したものである。音素認識結果は、この系列を隠れマルコフモデル (HMM) に入力して得る。

(5) 双対空間の利用

DNN は多段の MLP から成るため、学習に膨大な計算リソースを必要とする。そこで、この負荷を軽減するため、双対空間を通した特徴を MLP に加えることを検討した。ベクトル空間 V の双対空間 (Dual space) V^* は、汎関数 (functional; 関数の関数) を用いて表現される。ベクトル空間 V の直交基底を $\{\phi_i\}$ に採り、写像 $f_i \in V^*$ ($i=1, 2, \dots, m$) を $f_i(\phi_j) = \delta_{i,j} = 1$ ($i=j$)、もしくは 0 ($i \neq j$) と定めると、 V 上のベクトル v の線形写像は、 $f_i(v) = c_1 f_i(\phi_1) + \dots + c_i f_i(\phi_i) + \dots + c_m f_i(\phi_m) = c_i$ とスカラー値を定める。

ここでは、クラス k 毎の主成分分析 (PCA) から正規直交基底 $\{\phi^k_i\}$ を求め、更にこの基

底群からクラス共通の直交基底 $\{\phi_i\}$ を求める。学習データをクラス別 (提案手法 1) とクラス共通 (提案手法 2) の二種の双対空間で写像し、得られたベクトル (スカラー値を連結したものを) を MLP の入力とした際の結果を図 6 に示す。また、図 5 で説明した ΔT および ΔF を別々に双対空間で写像した後、MLP に通した結果 (提案手法 3) を同時に示した。

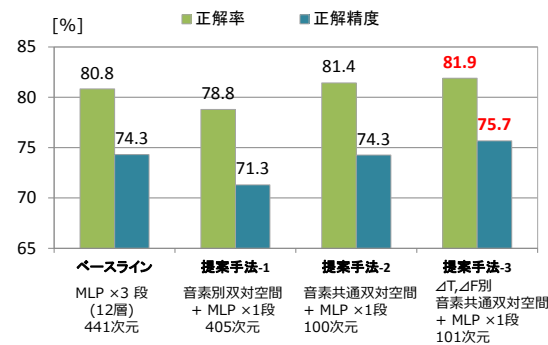


図 6 双対空間 + MLP (ニューラルネットワーク)

音素認識評価実験は、新聞記事読み上げコーパス (JNAS) を使用して行った (言語モデル無し、混合数 16)。図のベースラインは、MLP 3 段 (計 12 層) の結果を示す。局所特徴別 (ΔT , ΔF) に、音素共通双対空間で写像したベクトルデータを MLP 一段に通す方式 (提案手法-3) は、音素正解率、正解精度ともにベースラインを上回る結果を得ている。図の横軸最下段に次元数を示したが、この方式は演算量を大幅に軽減することができる。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計 8 件)

[1] Seng Kheang, Kouichi Katsurada, Yurie Iribe and Tsuneo Nitta: "Solving the phoneme conflict in Grapheme-To-Phoneme Conversion using a Two-Stage Neural Network-based approach", IEICE Trans. on Info. and System, 査読有, Vol. E97-D, No. 4, pp. 901-910 (2014).

[2] Narendyah W. Ariwardhani, Yurie Iribe, Kouichi Katsurada and Tsuneo Nitta: "Mapping Articulatory Features to Vocal-Tract Parameters for Voice Conversion", IEICE Trans. on Info. and System, 査読有, Vol. E97-D, No. 4, pp. 911-918 (2014).

[3] 桂田浩一, 勝浦広大, 入部百合絵, 新田恒雄: "Suffix Array を用いた高速音声検索語検出システムの性能評価", 電子情報通信学会論文誌, 査読有, Vol. J96-D, No. 10, pp. 2540-2548 (2013).

[4] 木村優志, 入部百合絵, 桂田浩一, 新田恒雄: "調音特徴一声道音響パラメータ変換を用いた調音特徴運動 HMM 音声合成", 電子情報通信学会論文誌, 査読有, Vol. J96-D,

No. 5, pp. 1356-1364 (2013).

[5] Yurie Iribe, Takuro Mori, Kouichi Katsurada and Tsuneo Nitta, "Generation of CG Animation Based on Articulatory Features for Pronunciation Training", The Journal of Information and Systems in Education, 査読有, Vol.11, No.1, pp.1-13 (2012).

[6] Narpendyah W. Ariwardhani, Masashi Kimura, Yurie Iribe, Kouichi Katsurada and Tsuneo Nitta, "Phoneme Recognition based on AF-HMMs with Optimal Parameter Set", Journal of Signal Processing, 査読有, Vol.16, No.6, pp.571-579 (2012).

[7] 木村優志, 澤田心大, 入部百合絵, 桂田浩一, 新田恒雄: "音声と画像シーンを用いた潜在意味解析に基づくタスク推定", 電気学会論文誌 C, 査読有, Vol.132, No.9, pp.1473-1480 (2012).

[8] Yurie Iribe, Hiroaki Nagaoka, Kouichi Katsurada and Tsuneo Nitta: "Web-Based Lecture System Using Slide Sharing for Classroom Questions and Answers", International Journal of Knowledge and Web Intelligence, 査読有, Vol.1, No.3/4, pp.243-255 (2010).

[学会発表] (国際会議 計 6/17 件,
国内学会発表 計 26/54 件;
学会発表 合計 32/71 件)

< 国際会議 >

[1] Kouichi Katsurada, Seiichi Miura, Kheang Seng, Yurie Iribe and Tsuneo Nitta, "Acceleration of Spoken Term Detection Using a Suffix Array by Assigning Optimal Threshold Values to Sub-Keywords", Proc. of InterSpeech 2013, pp.11-14 (2013 Lyon, France).

[2] Narpendyah Ariwardhani, Yurie Iribe, Kouichi Katsurada and Tsuneo Nitta, "Phoneme Recognition Based on AF-HMMs with an Optimal Parameter Set", Proc. of NCSP' 12, pp.170-173 (2012 Hawaii, USA).

[3] Tsuneo Nitta, Takayuki Onoda, Masashi Kimura, Yurie Iribe and Kouichi Katsurada: "Speech Synthesis based on Articulatory-Movement HMMs with Voice-source Codebooks", Proc. of Interspeech2011, pp.1841-1844 (2011 Florence, Italy).

[4] Kouichi Katsurada, Shinta Sawada, Shigeki Teshima, Yurie Iribe and Tsuneo Nitta, "Evaluation of Fast Spoken Term Detection Using a Suffix Array", Proc. of Interspeech2011, pp.909-912 (2011 Florence, Italy).

[5] Masashi Kimura, Takayuki Onoda, Yurie Iribe, Kouichi Katsurada and Tsuneo Nitta: "One-Model Speech Recognition and Synthesis Based on Articulatory Movement HMMs", Proc. of NCSP11, pp.392-395 (2011

Tianjin, China).

[6] Tsuneo Nitta, Takayuki Onoda, Masashi Kimura, Yurie Iribe and Kouichi Katsurada: "One-Model Speech Recognition and Synthesis Based on Articulatory Movement HMMs", Proc. of Interspeech2010, pp.2970-2973 (2010 Makuhari, Japan).

<国内学会発表>

[1] 桂田浩一, 入部百合絵, 新田恒雄, Suffix Array を用いた高速 STD におけるキーワード分割に関する理論的検討, 情報処理学会研究報告 SLP-89, No.16 (2011 芝浦工大, 東京).

[2] 三浦成一, 桂田浩一, 入部百合絵, 新田恒雄, Suffix Array を用いた高速 STD のための検索閾値の調整手法, 第8回音声ドキュメント処理ワークショップ (2014 豊橋市民センター, 豊橋市).

[3] 森本容介, 青木久美子, 桂田浩一, 石原元気, 三浦成一, 入部百合絵, 新田恒雄, Suffix Array を用いた音声検索語検出エンジンを利用する講義映像・音声の全文検索モジュールの開発, 電子情報通信学会技術研究報告, ET2013-124, pp.187-192 (2014 高知工業高専, 南国市).

[4] 三浦成一, 桂田浩一, 入部百合絵, 新田恒雄, Suffix Array を用いた高速音声検索語検出における検索閾値の調整法の改良, 日本音響学会秋季研究発表会講演論文集, 2-P-21 (2013 豊橋技科大, 豊橋市).

[5] Narpendyah Ariwardhani, Yurie Iribe, Kouichi Katsurada and Tsuneo Nitta, Mapping from Articulatory Movement to Vocal-Tract Spectrum for Voice Conversion, 日本音響学会春季研究発表会講演論文集, 3-P-6b (2013 東京工科大, 八王子市).

[6] 遠藤基, 入部百合絵, 桂田浩一, 新田恒雄, 調音運動音響モデルと長短音節言語モデルを用いた音素認識エンジン, 日本音響学会春季研究発表会講演論文集, 2-9-16 (2013 東京工科大, 八王子市).

[7] 佐々木俊, 入部百合絵, 桂田浩一, 新田恒雄: 双対空間における調音特徴抽出の検討, 日本音響学会春季研究発表会講演論文集, 2-9-15 (2013 東京工科大, 八王子市).

[8] 三浦成一, 桂田浩一, 入部百合絵, 新田恒雄, SuffixArray を用いた高速音声検索における検索閾値の付与法の改良, 第7回音声ドキュメント処理ワークショップ (2013 名古屋大, 名古屋市).

[9] 勝浦広大, 桂田浩一, 入部百合絵, 新田恒雄, 複数音声認識結果を用いて構築した Suffix Array に対する音声検索語検出, 情報処理学会研究報告 SLP-94, No.15, pp.1-6 (2012-12 東工大, 東京都).

[10] Kheang Seng, Kouichi Katsurada, Yurie Iribe, Tsuneo Nitta, Improving the performance of Letter-To-Phoneme conversion by using Two-Stage Neural Network, 情報処理学会研究報告 SLP-94, No.16, pp.1-7 (2012 東工大, 東京都).

[11] 三浦成一, 桂田浩一, 入部百合絵, 新田恒雄, Suffix Array を用いた高速音声検索における検索閾値の調整法の検討, 電気関係学会東海支部連合大会講演論文集, C3-6(2012 豊橋技科大, 豊橋市).

[12] 佐々木俊, 入部百合絵, 桂田浩一, 新田恒雄, 音素共通部分空間を結合重みとした調音運動 MLN に基づく音素認識, 電気関係学会東海支部連合大会講演論文集, C3-7 (2012 豊橋技科大, 豊橋市).

[13] 遠藤基, 桂田浩一, 入部百合絵, 新田恒雄, 調音特徴に基づく音素認識エンジンの提案, 電気関係学会東海支部連合大会講演論文集, C3-8 (2012 豊橋技科大, 豊橋市).

[14] Narpendyah Wisjnu Ariwardhani, Yurie Iribe, Kouichi Katsurada, Tsuneo Nitta: "Improvement of an AF-HMM based phoneme recognizer", 日本音響学会春季研究発表会講演論文集, 1-P-19 (2012 神奈川大, 横浜市).

[15] 勝浦広大, 桂田浩一, 入部百合絵, 森本容介, 辻靖彦, 青木久美子, 新田恒雄, 放送大学の講義音声を対象とした高速キーワード検索の性能評価, 日本音響学会春季研究発表会講演論文集, 3-7-11 (2012 神奈川大, 横浜市).

[16] 宮原伸也, 遠藤基, 入部百合絵, 桂田浩一, 新田恒雄, 調音特徴抽出における強調/抑制の効果, 日本音響学会春季研究発表会講演論文集, 1-7-3 (2012 神奈川大, 横浜市).

[17] Narpendyah Wisjnu Ariwardhani, Yurie Iribe, Kouichi Katsurada, Tsuneo Nitta, Phoneme Recognition based on AF-HMMs with Optimal State Configuration, 電子情報通信学会技術研究報告, SP-84, pp. 49-54 (2011 芝浦工大, 東京都).

[18] 池田大輔, 桂田浩一, 入部百合絵, 新田恒雄, 認識に使用する顔領域の違いによる読唇性能の比較, HAI シンポジウム 2012, 6-S (2011 京都工繊大, 京都市).

[19] Silasak Manosavanh, 入部百合絵, 桂田浩一, 林良子, 朱春躍, 新田恒雄, 音声から抽出した調音動作に基づく CG アニメーション生成, 日本音響学会秋季研究発表会講演論文集, 1-8-8 (2011 島根大, 松江市).

[20] 木村優志, 小野田高幸, 入部百合絵, 桂田浩一, 新田恒雄, 調音運動に基づくワンモデル音声認識合成方式, 電子情報通信学会技術研究報告, SP-41, pp. 1-6 (2011 定山溪, 札幌市).

[21] 新田恒雄, 小野田高幸, 荒木厚太, 入部百合絵, 桂田浩一, 調音運動の one-model を用いた音声認識・合成の改良, 人工知能学会全国大会論文集, 3B1-0S22c-2 (2011 岩手アイーナ, 盛岡市).

[22] 澤田心太, 桂田浩一, 入部百合絵, 新田恒雄, 音声ドキュメント高速検索におけるクエリ分割とマッチング手法の比較検討, 日本音響学会春季研究発表会講演論文集, 3-5-18 (2011 早稲田大, 東京都).

[23] 佐々木俊, 木村優志, 入部百合絵, 桂

田浩一, 新田恒雄, 調音特徴を用いた確率モデルに基づく VAD, 日本音響学会春季研究発表会講演論文集, 1-5-14 (2011 早稲田大, 東京都).

[24] 荒木厚太, 小野田高幸, 桂田浩一, 新田恒雄, 調音運動 HMM に基づく高精度音素認識の検討, 日本音響学会春季研究発表会講演論文集, 1-5-2 (2011 早稲田大, 東京都).

[25] 澤田心太, 桂田浩一, 入部百合絵, 新田恒雄, 高速音声ドキュメント検索における検索クエリ分割手法およびマッチング手法の比較評価, 第 5 回音声ドキュメント処理ワークショップ講演論文集, (2011 豊橋技科大, 豊橋市).

[26] 澤田心太, 桂田浩一, 新田恒雄, 入部百合絵, 手島茂樹, 大規模音声ドキュメントからの高速キーワード検索法の提案とその評価, 日本音響学会秋季研究発表会講演論文集, 2-9-10 (2010 関西大, 吹田市).

〔図書〕 (計 0 件)

〔産業財産権〕

○出願状況 (計 1 件)

名称者: 調音特徴抽出装置, 調音特徴抽出方法, 及び調音特徴抽出プログラム

発明者: 新田 恒雄

権利者: 国立大学法人豊橋技術科学大学

種類: 特許

番号: 2008-241072

出願年月日: 平成 20 年 9 月 19 日

国内外の別: 国内特許

○取得状況 (計 1 件)

名称: 音声検索装置および音声検索方法

発明者: 桂田 浩一, 手島 茂樹, 新田 恒雄

権利者: 国立大学法人豊橋技術科学大学

種類: 特許

番号: 第 5408631 号

取得年月日: 平成 25 年 11 月 15 日

国内外の別: 米国: 2014 年 1 月 7 日

(US8, 626, 508)

〔その他〕 なし

6. 研究組織

(1) 研究代表者

新田 恒雄 (Tsuneo Nitta)

早稲田大学グリーン・コンピューティング・システム研究機構 客員上級研究員・教授

(平成 24 年 3 月 31 日まで, 豊橋技術科学大学大学院知能情報工学専攻教授, 現在客員教授)

研究者番号: 70314101

(2) 研究分担者

桂田 浩一 (Kouichi Katsurada)

豊橋技術科学大学国際交流センター 准教授

研究者番号: 80324490

(3) 連携研究者

入部 百合絵 (Yurie Iribe)

愛知県立大科学情報科学部 助教

研究者番号: 40397500