

様 式 Z - 7

平成 2 6 年度科学研究費助成事業 実績報告書 (研究実績報告書)

1. 機関番号

1	4	6	0	3
---	---	---	---	---

 2. 研究機関名 奈良先端科学技術大学院大学
3. 研究種目名 基盤研究(B) 4. 研究期間 平成 2 6 年度 ~ 平成 2 8 年度
5. 課題番号

2	6	2	8	0	0	6	0
---	---	---	---	---	---	---	---
6. 研究課題名 統計的手法と生成過程モデリング手法の融合に基づく音声生成機能拡張技術の構築
7. 研究代表者

研究者番号	研究代表者名	所属部局名	職名
9 0 4 0 3 3 2 8	トダ トモキ	情報科学研究科	准教授
	戸田 智基		

8. 研究分担者

研究者番号	研究分担者名	所属研究機関名・部局名	職名
2 0 4 6 6 4 0 2	カメオカ ヒロカズ	日本電信電話株式会社 NTT コミュニケーション科学基礎研究所・メディア情報研究部	研究員
	亀岡 弘和		
3 0 2 6 3 4 2 9	ナカムラ サトシ	情報科学研究科	教授
	中村 哲		
3 0 3 2 4 9 7 4	サルワタリ ヒロシ	東京大学・情報理工学(系)研究科	教授
	猿渡 洋		
3 0 6 2 5 0 8 3	サクリアニ サクティ	情報科学研究科	助教
	サクリアニ サクティ		
7 0 6 3 3 4 2 8	ニュービッグ グラム	情報科学研究科	助教
	Neubig Graham		

9. 研究実績の概要

音声コミュニケーションにおいて物理的・身体的制約により生じている障壁を取り除くために、現存の音声生成過程において併用可能な音声変換基礎技術、および、音声生成機能を拡張する応用技術の構築を目指し、以下の課題に取り組んだ。

調音動作操作機能を備えた統計的声質変換技術の構築 (実施計画 1) : 統計的声質変換の基礎技術を拡張し、音声信号からの調音動作パラメータ推定処理、および、調音動作パラメータからの音声信号生成処理を実現し、これらの処理を繋ぎ合わせることで、調音動作操作機能を備えた統計的声質変換技術を構築した。また、調音動作操作時に適切な調音動作を保持するために、調音動作パラメータ補正技術を構築した。

音源生成器官動作操作機能を備えた統計的韻律変換技術の構築 (実施計画 2) : 音声F0パターン生成過程の確率モデルを考案し、音声信号からの音源生成器官動作指令推定技術を構築し、動作指令操作による音声F0パターン変換処理を実現した。さらに、言語情報からのF0パターン生成技術も構築した。

調音動作・音声同期収録データベースの構築 (実施計画 3) : 磁気センサを用いてリアルタイムに調音動作を収録できる発話観測システムを用いて、調音動作と音声信号の同期収録を実施した。男性話者 3 名を対象とし、各話者 50 ~ 100 文程度収録した。

音声生成機能を拡張する複数の応用技術の構築 (実施計画 4) : 統計的手法に基づく実時間声質変換基礎技術の性能を改善し、音声生成機能拡張技術として、発声障害者補助技術、外国語発声生成技術、体内伝導音声強調技術、ボーカルエフェクター / ボイスチェンジャー技術に対するベースラインシステムを構築した。

これらの研究成果をとりまとめ、国内外において多数の研究発表を行った。本研究成果は高い評価を受け、国内外において計 2 つの賞を受賞し、また、国際会議において 1 件の招待講演を実施するに至った。

10. キーワード

- | | | | |
|----------|----------|----------|----------|
| (1) 機能拡張 | (2) 音声合成 | (3) 音声変換 | (4) 信号処理 |
| (5) 統計処理 | (6) | (7) | (8) |

(注) ・印刷に当たっては、A 4 判 (縦長) ・両面印刷すること。

11. 現在までの達成度

(区分)(1) 当初の計画以上に進展している。

(理由)

当初の計画は全て実施した。さらに、音声生成機能を拡張する複数の応用技術の構築については、当初計画していなかった応用技術までも構築することができた。

12. 今後の研究の推進方策 等

(今後の推進方策)

当初の計画に基づき実施する。また、当初計画していなかった関連研究課題についても、積極的に取り組む。

(次年度使用額が生じた理由と使用計画)

(理由)

調音・音声データを1日で複数人収録することが可能となり、発話観測システムの使用回数を減らすことができたため。

(使用計画)

当初の計画を上回る研究成果が出ているため、国際会議での研究成果発表のための旅費に使用する。

13.研究発表(平成26年度の研究成果)

(雑誌論文) 計(11)件 うち査読付論文 計(11)件

著者名		論文標題			
Kou Tanaka, Tomoki Toda, Graham Neubig, Sakriani Sakti, Satoshi Nakamura		Direct F0 control of an electrolarynx based on statistical excitation feature prediction and its evaluation through simulation			
雑誌名	査読の有無	巻	発行年	最初と最後の頁	
Proceedings of INTERSPEECH	有		2 0 1 4	31 ~ 35	
掲載論文のDOI(デジタルオブジェクト識別子)					
なし					

著者名		論文標題			
Kento Kadowaki, Tatsuma Ishihara, Nobukatsu Hojo, Hirokazu Kameoka		Speech prosody generation for text-to-speech synthesis based on generative model of F0 contours			
雑誌名	査読の有無	巻	発行年	最初と最後の頁	
Proceedings of INTERSPEECH	有		2 0 1 4	2322 ~ 2326	
掲載論文のDOI(デジタルオブジェクト識別子)					
なし					

著者名		論文標題			
Kazuhiro Kobayashi, Tomoki Toda, Graham Neubig, Sakriani Sakti, Satoshi Nakamura		Statistical singing voice conversion with direct waveform modification based on the spectrum differential			
雑誌名	査読の有無	巻	発行年	最初と最後の頁	
Proceedings of INTERSPEECH	有		2 0 1 4	2514 ~ 2518	
掲載論文のDOI(デジタルオブジェクト識別子)					
なし					

著者名		論文標題			
Sakura Tsuruta, Kou Tanaka, Tomoki Toda, Graham Neubig, Sakriani Sakti, Satoshi Nakamura		An evaluation of target speech for a nonaudible murmur enhancement system in noisy environments			
雑誌名	査読の有無	巻	発行年	最初と最後の頁	
Proceedings of APSIPA ASC	有		2 0 1 4	1~4	
掲載論文のDOI(デジタルオブジェクト識別子)					
10.1109/APSIPA.2014.7041618					

著者名		論文標題			
Kazuhiro Kobayashi, Tomoki Toda, Tomoyasu Nakano, Masataka Goto, Graham Neubig, Sakriani Sakti, Satoshi Nakamura		Gender-dependent spectrum differential models for perceived age control based on direct waveform modification in singing voice conversion			
雑誌名	査読の有無	巻	発行年	最初と最後の頁	
Proceedings of APSIPA ASC	有		2 0 1 4	1~4	
掲載論文のDOI(デジタルオブジェクト識別子)					
10.1109/APSIPA.2014.7041590					

著者名		論文標題			
Kou Tanaka, Tomoki Toda, Graham Neubig, Sakriani Sakti, Satoshi Nakamura		An inter-speaker evaluation through simulation of electrolarynx control based on statistical F0 prediction			
雑誌名	査読の有無	巻	発行年	最初と最後の頁	
Proceedings of APSIPA ASC	有		2 0 1 4	1~4	
掲載論文のDOI(デジタルオブジェクト識別子)					
10.1109/APSIPA.2014.7041593					

著者名		論文標題			
Shinnosuke Takamichi, Tomoki Toda, Alan W. Black, Satoshi Nakamura		Modulation spectrum-based post-filter for GMM-based voice conversion			
雑誌名	査読の有無	巻	発行年	最初と最後の頁	
Proceedings of APSIPA ASC	有		2 0 1 4	1 ~ 4	
掲載論文のDOI(デジタルオブジェクト識別子)					
10.1109/APSIPA.2014.7041540					

著者名		論文標題			
Shinnosuke Takamichi, Tomoki Toda, Alan W. Black, Satoshi Nakamura		Modified post-filter to recover modulation spectrum for HMM-based speech synthesis			
雑誌名	査読の有無	巻	発行年	最初と最後の頁	
Proceedings of GlobalSIP	有		2 0 1 4	547 ~ 551	
掲載論文のDOI(デジタルオブジェクト識別子)					
10.1109/GlobalSIP.2014.7032177					

著者名		論文標題			
Takuya Higuchi, Hirokazu Kameoka		Unified approach for underdetermined BSS, VAD, dereverberation and DOA estimation with multichannel factorial HMM			
雑誌名	査読の有無	巻	発行年	最初と最後の頁	
Proceedings of GlobalSIP	有		2 0 1 4	562 ~ 566	
掲載論文のDOI(デジタルオブジェクト識別子)					
10.1109/GlobalSIP.2014.7032180					

著者名		論文標題			
Tomoki Toda		Augmented speech production based on real-time statistical voice conversion			
雑誌名	査読の有無	巻	発行年	最初と最後の頁	
Proceedings of GlobalSIP	有		2 0 1 4	592 ~ 596	
掲載論文のDOI(デジタルオブジェクト識別子)					
10.1109/GlobalSIP.2014.7032186					

著者名		論文標題			
Ryosuke Sugiura, Yutaka Kamamoto, Noboru Harada, Hirokazu Kameoka, Takehiro Moriya		Golomb-Rice coding optimized via LPC for frequency domain audio coder			
雑誌名	査読の有無	巻	発行年	最初と最後の頁	
Proceedings of GlobalSIP	有		2 0 1 4	1024 ~ 1028	
掲載論文のDOI(デジタルオブジェクト識別子)					
10.1109/GlobalSIP.2014.7032276					

(学会発表) 計(28)件 うち招待講演 計(0)件

発表者名		発表標題	
門脇 健人, 北条 伸克, 石原 達馬, 亀岡 弘和		音声F0パターン生成過程の確率モデルによるテキストからの韻律生成	
学会等名	発表年月日	発表場所	
電子情報通信学会 / 日本音響学会 音声研究会	2014年05月24日 ~ 2014年05月25日	日本大学文理学部キャンパス(東京都世田谷区)	

発表者名		発表標題	
田中 宏, 戸田 智基, Graham Neubig, Sakriani Sakti, 中村 哲		統計的音源予測に基づく電気式人工喉頭制御法のシミュレーションによる評価	
学会等名	発表年月日	発表場所	
電子情報通信学会 / 日本音響学会 音声研究会	2014年06月19日 ~ 2014年06月20日	近江町交流プラザ(石川県金沢市)	

発表者名	発表標題	
田中 宏, 戸田 智基, Graham Neubig, Sakriani Sakti, 中村 哲	統計的音源予測を用いた電気式人工喉頭制御法の評価	
学会等名	発表年月日	発表場所
日本音響学会2014年秋季研究発表会	2014年09月03日～2014年09月05日	北海学園大学豊平キャンパス(北海道札幌市)

発表者名	発表標題	
鶴田 さくら, 田中 宏, 戸田 智基, Graham Neubig, Sakriani Sakti, 中村 哲	雑音環境下での受聴を想定した非可聴つぶやき強調処理における変換音声有声化の効果	
学会等名	発表年月日	発表場所
日本音響学会2014年秋季研究発表会	2014年09月03日～2014年09月05日	北海学園大学豊平キャンパス(北海道札幌市)

発表者名	発表標題	
小林 和弘, 戸田 智基, 中野 倫靖, 後藤 真孝, Graham Neubig, Sakriani Sakti, 中村 哲	性別依存重回帰混合正規分布モデルに基づく差分スペクトル補正による歌声の知覚年齢制御法	
学会等名	発表年月日	発表場所
日本音響学会2014年秋季研究発表会	2014年09月03日～2014年09月05日	北海学園大学豊平キャンパス(北海道札幌市)

発表者名	発表標題	
門脇 健人, 亀岡 弘和	音声F0パターン生成過程の確率モデルによるテキストからの韻律生成及びその評価	
学会等名	発表年月日	発表場所
日本音響学会2014年秋季研究発表会	2014年09月03日～2014年09月05日	北海学園大学豊平キャンパス(北海道札幌市)

発表者名	発表標題	
久保 和隆, 小林 和弘, 戸田 智基, Graham Neubig, Sakriani Sakti, 中村 哲	重回帰混合正規分布モデルに基づく声質制御における制御パラメータの設計	
学会等名	発表年月日	発表場所
電子情報通信学会 / 日本音響学会 音声研究会	2014年11月13日 ~ 2014年11月14日	九州大学筑紫キャンパス(福岡県春日市)

発表者名	発表標題	
鶴田 さくら, 田中 宏, 戸田 智基, Graham Neubig, Sakriani sakti, 中村 哲	雑音環境下における明瞭性に着目した非可聴つぶやき強調処理における目標音声の評価	
学会等名	発表年月日	発表場所
電子情報通信学会 / 日本音響学会 音声研究会	2014年11月13日 ~ 2014年11月14日	九州大学筑紫キャンパス(福岡県春日市)

発表者名	発表標題	
小林 和弘, 戸田 智基, 中野 倫靖, 後藤 真孝, Graham Neubig, Sakriani Sakti, 中村 哲	知覚年齢をリアルタイムに制御可能な歌声声質制御インタフェース	
学会等名	発表年月日	発表場所
第22回インタラクティブシステムとソフトウェアに関するワークショップ (WISS 2014)	2014年11月26日 ~ 2014年11月28日	浜名湖ロイヤルホテル(静岡県浜松市)

発表者名	発表標題	
Patrick Lumban Tobing, Tomoki Toda, Graham Neubig, Sakriani Sakti, Satoshi Nakamura, Ayu Purwarianti	Articulatory controllable speech modification using sequential inversion and production mapping with Gaussian mixture models	
学会等名	発表年月日	発表場所
電子情報通信学会 / 日本音響学会 音声研究会	2014年12月15日 ~ 2014年12月16日	東京工業大学すずかけ台キャンパス(神奈川県横浜市)

発表者名	発表標題	
大島 悠司, 高道 慎之介, 戸田 智基, Graham Neubig, Sakriani Sakti, 中村 哲	HMMを用いた日本人英語音声合成における話者性を保持した韻律補正	
学会等名	発表年月日	発表場所
電子情報通信学会 / 日本音響学会 音声研究会	2014年12月15日 ~ 2014年12月16日	東京工業大学すずかけ台キャンパス(神奈川県横浜市)

発表者名	発表標題	
西垣 友理, 高道 慎之介, 戸田 智基, Graham Neubig, Sakriani Sakti, 中村 哲	音声入力に基づく韻律制御機能を有するHMM音声合成システム	
学会等名	発表年月日	発表場所
電子情報通信学会 / 日本音響学会 音声研究会	2014年12月15日 ~ 2014年12月16日	東京工業大学すずかけ台キャンパス(神奈川県横浜市)

発表者名	発表標題	
高道 慎之介, 戸田 智基, Alan W. Black, 中村 哲	統計的パラメトリック音声合成のための変調スペクトルを考慮した音声パラメータ生成アルゴリズム	
学会等名	発表年月日	発表場所
情報処理学会 音声言語情報処理研究会	2015年02月27日 ~ 2015年02月28日	合歡の郷ホテル&リゾート(三重県志摩市)

発表者名	発表標題	
大島 悠司, 高道 慎之介, 戸田 智基, Sakriani Sakti, Graham Neubig, 中村 哲	韻律・音韻の部分補正に基づく話者性を保持した日本人英語音声合成と英語習熟度が与える影響	
学会等名	発表年月日	発表場所
情報処理学会 音声言語情報処理研究会	2015年02月27日 ~ 2015年02月28日	合歡の郷ホテル&リゾート(三重県志摩市)

発表者名	発表標題	
高道 慎之介, 戸田 智基, Alan W. Black, 中村 哲	統計的パラメトリック音声合成のための変調スペクトル制約付きトラジェクトリ学習アルゴリズム	
学会等名	発表年月日	発表場所
電子情報通信学会 / 日本音響学会 音声研究会	2015年03月02日 ~ 2015年03月03日	南の美ら花ホテルミヤヒラ (沖縄県石垣市)

発表者名	発表標題	
樋口卓哉, 亀岡弘和	多チャンネル階乗隠れマルコフモデルによる音響情景分析のための統合的アプローチ	
学会等名	発表年月日	発表場所
電子情報通信学会 / 日本音響学会 音声研究会	2015年03月02日 ~ 2015年03月03日	南の美ら花ホテルミヤヒラ (沖縄県石垣市)

発表者名	発表標題	
田中 宏, 戸田 智基, Graham Neubig, Sakriani Sakti, 中村 哲	統計的音源予測に基づく電気式人工喉頭のリアルタイム制御	
学会等名	発表年月日	発表場所
日本音響学会2015年春季研究発表会	2015年03月16日 ~ 2015年03月18日	中央大学後楽園キャンパス (東京都文京区)

発表者名	発表標題	
俱羅 真也, 高道 慎之介, 戸田 智基, 亀岡 弘和, Graham Neubig, Sakriani Sakti, 中村 哲	非母語話者の音声に対する継続長補正の評価	
学会等名	発表年月日	発表場所
日本音響学会2015年春季研究発表会	2015年03月16日 ~ 2015年03月18日	中央大学後楽園キャンパス (東京都文京区)

発表者名	発表標題	
大島 悠司, 高道 慎之介, 戸田 智基, Graham Neubig, Sakriani Sakti, 中村 哲	日本人英語音声合成における話者性を保持した韻律補正法と英語習熟度が与える影響	
学会等名	発表年月日	発表場所
日本音響学会2015年春季研究発表会	2015年03月16日～2015年03月18日	中央大学後楽園キャンパス(東京都文京区)

発表者名	発表標題	
高道 慎之介, 戸田 智基, Alan W Black, 中村 哲	統計的パラメトリック音声合成における変調スペクトルを考慮したパラメータ生成法	
学会等名	発表年月日	発表場所
日本音響学会2015年春季研究発表会	2015年03月16日～2015年03月18日	中央大学後楽園キャンパス(東京都文京区)

発表者名	発表標題	
高道 慎之介, 戸田 智基, Alan W Black, 中村 哲	統計的パラメトリック音声合成における変調スペクトル制約付きトラジェクトリ学習	
学会等名	発表年月日	発表場所
日本音響学会2015年春季研究発表会	2015年03月16日～2015年03月18日	中央大学後楽園キャンパス(東京都文京区)

発表者名	発表標題	
西垣 友理, 高道 慎之介, 戸田 智基, Graham Neubig, Sakriani Sakti, 中村 哲	音声入力による韻律制御機能を有するHMM音声合成システムの改良	
学会等名	発表年月日	発表場所
日本音響学会2015年春季研究発表会	2015年03月16日～2015年03月18日	中央大学後楽園キャンパス(東京都文京区)

発表者名	発表標題	
久保 和隆, 小林 和弘, 戸田 智基, Graham Neubig, Sakriani Sakti, 中村 哲	重回帰混合正規分布モデルに基づく声質制御における精度改善	
学会等名	発表年月日	発表場所
日本音響学会2015年春季研究発表会	2015年03月16日～2015年03月18日	中央大学後楽園キャンパス(東京都文京区)

発表者名	発表標題	
Patrick Lumban Tobing, Kazuhiro Kobayashi, Tomoki Toda, Graham Neubig, Sakriani Sakti, Satoshi Nakamura	Articulatory controllable speech modification based on gaussian mixture models with direct waveform modification using spectrum differential	
学会等名	発表年月日	発表場所
日本音響学会2015年春季研究発表会	2015年03月16日～2015年03月18日	中央大学後楽園キャンパス(東京都文京区)

発表者名	発表標題	
鶴田 さくら, 田中 宏, 戸田 智基, Graham Neubig, Sakriani Sakti, 中村 哲	非可聴つばやき強調音声の雑音環境下における明瞭性改善に関する検討	
学会等名	発表年月日	発表場所
日本音響学会2015年春季研究発表会	2015年03月16日～2015年03月18日	中央大学後楽園キャンパス(東京都文京区)

発表者名	発表標題	
田尻 祐介, 鶴田 さくら, 田中 宏, 戸田 智基, Graham Neubig, Sakriani Sakti, 中村 哲	空気 / 体内伝導マイクを併用した雑音環境下における非可聴つばやき強調	
学会等名	発表年月日	発表場所
日本音響学会2015年春季研究発表会	2015年03月16日～2015年03月18日	中央大学後楽園キャンパス(東京都文京区)

発表者名	発表標題	
小林 和弘, 戸田 智基, Graham Neubig, Sakriani Sakti, 中村 哲	差分スペクトル補正に基づく歌声声質変換におけるパラメータ生成法に関する調査	
学会等名	発表年月日	発表場所
日本音響学会2015年春季研究発表会	2015年03月16日～2015年03月18日	中央大学後楽園キャンパス(東京都文京区)

発表者名	発表標題	
門脇 健人, 北条 伸克, 亀岡 弘和	複合ウェーブレットモデルとF0パターン生成過程の確率モデルを用いたテキスト音声合成	
学会等名	発表年月日	発表場所
日本音響学会2015年春季研究発表会	2015年03月16日～2015年03月18日	中央大学後楽園キャンパス(東京都文京区)

(図書) 計(0)件

著者名	出版社		
書名		発行年	総ページ数

14. 研究成果による産業財産権の出願・取得状況

(出願) 計(1)件

産業財産権の名称	発明者	権利者	産業財産権の種類、番号	出願年月日	国内・外国の別
電気式人工喉頭装置	戸田 智基, 田中 宏, 他3名	同左	特許、PCT/JP2014/069274	2014年07月22日	外国

(取得) 計(0)件

産業財産権の名称	発明者	権利者	産業財産権の種類、番号	取得年月日	国内・外国の別
				出願年月日	

15.備考

--