

## 科学研究費助成事業 研究成果報告書

平成 27 年 6 月 15 日現在

機関番号：12601  
研究種目：基盤研究(B)  
研究期間：2012～2014  
課題番号：24300068  
研究課題名(和文)基本周波数パターン生成過程モデルによる統計モデリング音声合成の韻律制御の高度化  
  
研究課題名(英文)Advanced method of prosody control in statistical-based speech synthesis using generation process model of fundamental frequency contours  
  
研究代表者  
広瀬 啓吉(Hirose, Keikichi)  
  
東京大学・情報理工学(系)研究科・教授  
  
研究者番号：50111472  
  
交付決定額(研究期間全体)：(直接経費) 13,700,000円

研究成果の概要(和文)：統計的音声合成の学習と合成において、基本周波数パターン(F0)生成過程モデルの制約を適用することで、柔軟な韻律制御と音質向上を実現することを目的として研究を進め、モデルで近似したF0パターンをHMMの学習に用いる手法等を開発するとともに、指令の差分に着目した焦点制御等を実現した。特に、F0パターンをモデルにより階層表現することで、言語情報との明確な対応を保った良好な韻律制御を実現した。この他、行列変量混合ガウス分布による話者の特徴表現、話者依存サブネットワークを有するDeep Neural Networkによる複数話者変換を達成した。中国語についても研究を進め、プロトタイプ音声翻訳を行った。

研究成果の概要(英文)：Research works were conducted with the aim of realizing flexible control of prosody and better speech quality in statistical-based speech synthesis by applying constraints of the generation process model of fundamental frequency (F0) contours. Several methods were developed including one to use F0 contours approximated by the model for HMM training. In the method, hierarchical F0 contours based on the model were treated separately by the multi-stream scheme, leading to a better prosody control keeping clear relations with linguistic information. Lexical emphasis was realized by manipulating the model commands (prosody conversion). Better speaker conversions were realized in multi-speaker case through matrix-variate Gaussian mixture model and deep neural network with speaker-dependent sub-networks. Research works were conducted also for Chinese, with preliminary experiments on speech translation.

研究分野：音声言語情報処理

キーワード：基本周波数パターン 生成過程モデル 統計的音声合成 韻律制御 音声変換 談話の焦点 マルチストリーム学習 行列変量GMM

### 1. 研究開始当初の背景

社会の情報化の進展に伴い、音声を紹介し、使いやすい機械-人間インターフェースに対するニーズが高まっている。また、国際化の進展に伴い、言語の違いを意識することなく円滑な音声コミュニケーションを行い得る音声自動翻訳システム、あるいは、計算機を援用した効果的な外国語発音教育システムの確立が期待されている。これらを実現するためには、要素技術としての音声認識・合成の高度化が必須であり、精力的な研究が各所で進められてきた。音声合成についてみると、波形選択接続手法の確立により合成音声の品質は格段に向上したが、そこで得られる音声の声質や調子は、あらかじめ用意された音声コーパスのそれに限られる。単に、テキストから読み上げ調子を合成するだけでは、不十分であって、様々な声質や調子を自由に実現し得る音声合成手法、すなわち少量の音声コーパスによって(理想は音声コーパスなしに)、任意の声質・調子の音声を合成し得る“柔軟な音声合成”の開発が求められている。この様な観点から、統計的な枠組みでの分析合成に基づく音声合成、特に、隠れマルコフモデル(HMM)に基づく音声合成が注目され、音素モデルを適応することにより、少量の音声コーパスから種々の声質・調子の音声を得られるようになってきている。当初は、音質に問題があったが、音声の声道伝達特性と音源特性の分離を安定して行い得る音声分析手法 STRAIGHT の導入などによって、品質が向上している。しかしながら、韻律に着目すると、それを表現する主要な音響的特徴である基本周波数( $F_0$ )を単に各フレームでの値として取り扱っているため、生成される韻律が不自然なものとなることが多い。時間差分値( $\Delta$ 特徴量)あるいは長単位での  $F_0$  変動量(Global Variance)を導入することによって、これを解消しようとする試みが行われているが、そもそも  $F_0$  の時間変化パターン( $F_0$ パターン)に代表される韻律は、音節、単語、句、さらには文といった長時間に渡るもので、フレーム毎の表現ではおのずと限界がある。さらに、音声合成で生成される  $F_0$  パターンについて、その背後にある言語情報、あるいはパラ/非言語情報との明示的な関係が得られないため、得られた合成音声の韻律を変換してこれらの情報を追加的に付与するといった柔軟な操作が困難となる。

声質や調子の変換としては、MLLR などの HMM 適応の他、言語内容が同一の音声から得られるパラレルコーパスを用いて混合ガウス分布モデル(GMM)を構築することが話者変換等で一般的に行われている。変換対象音声に直接存在しない場合にも変換可能であり、発話者の話者性を保った音声翻訳といった試みも行われている。しかしながら、フレーム毎の  $F_0$  値をそのまま用いた変換では、適正な変換を行うことが困難であり、発話全体に対する  $F_0$  の平均と分散を用いた線

形的な変換が行われるに留まっている[1]。

これに対し、本研究では、基本周波数パターン生成過程のモデル(生成過程モデル[2]、図1)による韻律の表現を導入することによって、HMM 音声合成や GMM による変換における韻律制御の問題点を解決し、韻律の観点から、音声合成の高品質化と柔軟性向上を達成することを目指す。

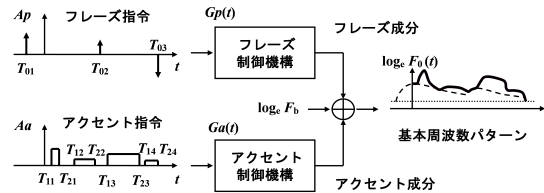


図1. 生成過程モデル

### 2. 研究の目的

代表者らは、テキストから生成過程モデルの枠組みで韻律を生成するコーパスベース手法を開発し、HMM 音声合成と統合することにより音質の向上が得られ、また、生成過程モデルの指令の差分に着目することで、談話の焦点を付与した強調音声が、少量の音声コーパスから得られることを示しているが[3]、分節的特徴と韻律的特徴の一体制御という HMM 音声合成の利点が失われる結果ともなっていた。本研究では、HMM 音声合成の過程に生成過程モデルの枠組みをより積極的に取り入れることを目指す。具体的には、生成過程モデルの制約を、音声の合成時のみならず、HMM の学習時にも導入する。生成過程モデルの導入によって、長時間に渡る韻律制御を適切に行うことが可能となり、また、合成の結果得られる韻律と、言語情報、パラ/非言語情報との明確な対応が得られ、少量の音声コーパスからでも種々の声質・調子の変換が良好に達成される。本研究の結果、高品質かつ柔軟な韻律制御の音声合成が達成され、ユーザの好みや状況に応じて適応的に変化する音声出力が可能でユーザフレンドリな音声対話、元の音声に含まれる声質・調子を保持した音声翻訳など、機械-人間システムの飛躍的な向上が期待される。さらに、種々に韻律を変化させた合成音声によって、非母語音声の特徴の詳細なモデル化が可能となり、発音教育への寄与も大きい。

### 3. 研究の方法

本研究では、学習と合成の両フェーズで HMM 音声合成の高品質化を目指す。HMM の学習のフェーズでは、学習用音声コーパスの有声/無声の判定誤りを含む  $F_0$  抽出誤りにより、HMM の状態の有声/無声の割り当て誤りが生じ、合成音声の品質劣化を招く。これに対し、生成過程モデルで近似することで  $F_0$  パターンの修正・補間を行うことにより、HMM の状態を有声/無声に明示的に割り当てることなく合成を行うことを可能とする。な

お、観測される F0 パターンには、マイクロプロソディーなど、生成過程モデルで表現されない F0 の動きが含まれる。この様な F0 残差を HMM 音声合成において個別のストリームとして取り扱うことを行う。さらに、生成過程モデルでは、F0 パターンがフレーズ成分にアクセント成分が重畳したものと表現されるが、両者は言語情報等との対応の様子が異なり、HMM 音声合成においても別個のストリームとして取り扱うことで、よりよい表現（コンテキストクラスタリング）が行えると考えられる。結局、F0 パターンを、フレーズ成分、アクセント成分、F0 残差の 3 ストリームとして処理する事を行う。合成のフェーズでの生成過程モデルの制約としては、既に HMM 音声合成で生成される F0 パターンに対し、生成過程モデルによる逐次近似を行なって修正し、合成に用いることで音質の向上が得られることを示しているが[4]、HMM 音声合成の過程で直接制約をかける方策を探る。以上の結果、HMM 音声合成で得られる F0 パターンと生成過程モデルの指令との良好な対応が得られ、柔軟な韻律制御が可能となる。また、我々は、生成過程モデルの指令の差分に着目することで韻律変換を行う手法を開発しており[5]、これにより、種々の声質・調子の変換を達成する。平静音声から感情音声への変換など、変換の前後で韻律の構造が大きく異なることが考えられ、特に本研究では指令が 1 対 1 に対応しないときの対処手法を探る。また、生成過程モデル指令の差分による手法では、同一の文の（パラレル）音声コーパスを前提としているが、平均的な韻律からのずれを生成過程モデルで表現することにより、この前提によらない手法の開発も進める。これによって発話者の話者性を保った音声翻訳（言語変換）を達成することも視野に入れる。本研究では、日本語の他、声調言語として複雑な F0 パターンを有する中国語についても研究を進め、開発した手法の頑健性を検証するとともに、音声変換手法の検証のために両言語の音声翻訳も試みる。

#### <引用文献>

1. A. Kain and M. Macon, "Spectral voice conversion for text-to-speech synthesis." *Proc. IEEE ICASSP*, Vol. 1, pp.285-288 (2002).
2. H. Fujisaki and K. Hirose, "Analysis of voice fundamental frequency contours for declarative sentences of Japanese," *J. Acoust. Soc. Japan (E)*, Vol.5, No.4, pp.233-242 (1984-10).
3. K. Ochi, K. Hirose, and N. Minematsu, "Control of prosodic focus in corpus-based generation of fundamental frequency contours of Japanese based on the generation process model," *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing*, Taipei, pp.4485-4488 (2009-4).
4. T. Matsuda, K. Hirose and N. Minematsu, "HMM-based synthesis of fundamental

frequency contours using the generation process model," *Journal of Signal Processing*, vol.14, no.4, pp.277-280 (2010-7).

5. K. Hirose, K. Ochi, R. Mihara, H. Hashimoto, D. Saito, and N. Minematsu, "Adaptation of prosody in speech synthesis by changing command values of the generation process model of fundamental frequency," *Proc. INTERSPEECH*, Florence, pp.2793-2796 (2011-8).

#### 4. 研究成果

- (1) 生成過程モデルを用いた階層表現による HMM 音声合成の F0 制御

まず、学習コーパスの F0 パターンを生成過程モデルで近似したものに置き換えて音素 HMM を学習して HMM 音声合成を行った。その結果、従来の HMM 音声合成と比較し、高品質が得られることを聴取実験により示した。この方法では、マイクロプロソディーなど生成過程モデルで表現されない F0 の動きは音声合成に反映されない。音質の観点からは、必ずしも大きな問題とはならないが、本研究では、これを F0 残差として取り扱うことを行った。さらに生成過程モデルのフレーズ成分は統語構造など広い範囲の言語情報との関連が深いのに対し、アクセント成分はアクセント型などより狭い範囲の言語情報との関係が深い。また F0 残差は音素の情報などさらに狭い範囲の情報との関係が深いと考えられ、これらを 1 つの F0 ストリームとして HMM 合成で取り扱うよりも、個別の 3 ストリームとして取り扱う方が、言語情報とのよりよい対応が取れ、合成音声の品質の向上に結び付く可能性がある。この様な考察から、F0 をフレーズ成分、アクセント成分、F0 残差の 3 ストリームとして階層表現し、HMM の学習と音声合成を行った。なお、各ストリームを時間的に連続したものとして取り扱い、従来の多空間確率分布 HMM でなく、連続 F0HMM とすることで、有声・無声の切り替え等での品質劣化に対処する事を行った。(F0 残差は無声ではゼロとした。)その結果、3 ストリーム化しない従来の HMM 音声合成と比較し、自然音声の F0 パターンに近い F0 パターンが生成されることを示した。(対数 F0 の差分が 75%に低下)

フレーズ成分の初めのアクセント成分は大きいなど、アクセント成分の生成には、フレーズ成分がどの様に生成されたかが大きく関わってくるが、この様な関係は、3 ストリームでは表現されない。そこで、フレーズ成分、アクセント成分、F0 残差を 3 次元のベクトルとして表現することも行った。対数 F0 の差分はわずかに低下したが、顕著な効果は認められなかった。フレーズ成分とアクセント成分の関係をどの様に表現するかについては、さらに今後検討が必要である。

対数 F0 の差分といった客観評価に加え、合成音声の聴取による主観評価も行った。従

来の HMM 音声合成と比較し、3 ストリームのほうが、若干、高い評価を得たが、有意差は認められなかった。

3 ストリームとする利点は、図 2 のコンテキストクラスタリングの結果にも示されている。フレーズ成分では、呼気段落など長い時間単位に関する質問項目が root node の近くに見られるのに対し、アクセント成分では、アクセント句、F0 残差では、音素に関する質問項目が見られる。

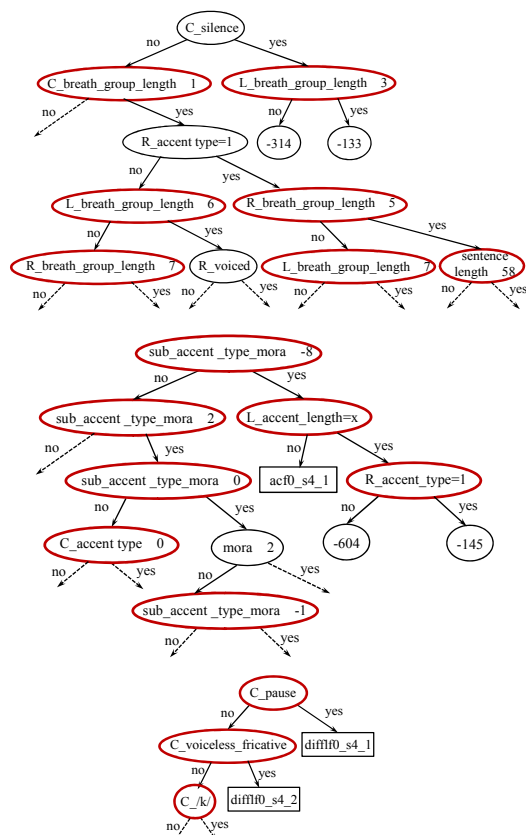


図 2. 各ストリームに対するコンテキストクラスタリングの結果 (root node 近く)。上から下へ、フレーズ成分、アクセント成分、F0 残差の結果。赤で示した質問項目がそれぞれ、呼気段落長 / 文長、アクセント句 (型)、音素に関するもの。

フレーズ成分、アクセント成分、F0 残差を個別に取り扱うことにより、生成された F0 パターンについて、生成過程モデルの当てはめを自動かつ高精度に行うことが可能であることを確認した。これは、指令の差分に着目した韻律変換の前提となる。

(2) 生成過程モデルにより近似された F0 パターンを用いた音声合成と F0 モデルの指令の差分に着目した韻律の適応

従来の HMM 音声合成で得られる F0 パターンを生成過程モデルで近似することで、音質が向上することを、合成音声の聴取実験により確認した。なお、モデル近似は、我々が開発した指令の自動抽出手法によって行った。

これは母音部の F0 パターンのみからフレーズ指令 / 成分を推定し、その後、アクセント指令 / 成分を推定するものであり、各モーラについて、言語情報から得られる高 F0 レベル・低 F0 レベル情報を用い、フレーズ指令 / アクセント指令の位置と大きさを推定する。

この HMM 音声合成の学習に用いた音声は、特に談話の焦点を考慮したものではない。そのため、合成される音声も特に焦点を考慮しないものとなる。これに対し、指定した語句に焦点を置いて発声した音声と特に置かないで発声した音声の平行コーパスから、焦点による指令の大きさの差分を学習し、それによってモデル近似した F0 パターンを変更することにより、指定した語句に焦点を置いた音声とすることが可能なことを示した。さらに、中国語についてもこの手法が有効であることを確認した。

(3) F0 残差による学習データの選択

学習コーパスの分析精度は、HMM 音声合成の品質に大きく影響するが、F0 抽出に誤りがある時点では、分析結果全体も誤りである可能性が高い。F0 パターン生成過程モデルと比較した差が大きい時点は分析結果に問題があると考え、学習から除外することを行った。音素単位で学習から除外することで、合成音声の品質が向上することを示した。明示的に除外せずにコンテキストクラスタリングの過程に反映させる方法も開発し、同様の効果が得られた。

(4) 声調核モデルによる中国語韻律の制御

声調言語である中国語は、日本語等と比べ、F0 の動きが複雑であり、生成過程モデルによる近似を自動的に行うことが困難である。このため、学習に十分なコーパスを用意することが難しい。一方、中国語は、音節単位での F0 の動きを表現するのに適している。この様な観点から、音節の中央部で F0 の動きが安定した部分 (声調核) の F0 パターンのみに着目し、その他は遷移部分として線形接続する声調核モデルに基づく音声合成手法を開発し、合成音声の聴取実験によりその有効性を示した。さらに、平静音声から感情音声への音声変換も小さなコーパスで良好に行い示した。中国語と同じく声調言語であるタイ語についても、実験を行い、声調核モデルの有効性を示した。

(5) 多人数間の話者変換手法

多人数話者の音声データを効率よく利用し、話者変換精度を上げる手法として、各話者の特徴を行列変量混合ガウス分布として表現した上で、変換モデルを構築する手法を開発した。これにより、従来の混合ガウス分布による固有声質変換法を超える性能を達成した。また、Deep Neural Network に基づく多人数話者間の声質変換手法を開発した。

1 つの話者非依存サブネットワークと複数話者の話者依存サブネットワークからなる構造とすることで、話者非依存な特徴量変換と話者依存の特徴量変換とを分けて効率的な学習を実現し、従来手法を超える変換性能を達成した。

#### (6) 音声翻訳の実験

日本語 中国語のプロトタイプ音声翻訳システム構築し、話者性を保存した言語変換を試みることで、これまでの成果の検証を行った。

#### 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6件)

1. Ya Li, Jianhua Tao, Keikichi Hirose, Xiaoying Xu, and Wei Lai, "Hierarchical stress modeling and generation in mandarin for expressive Text-to-Speech," *Speech Communication*, Vol.72, pp.59-73 (2015-5). 査読有  
DOI: 10.1016/j.specom.2015.05.003
2. 橋本浩弥, 齋藤大輔, 峯松信明, 広瀬啓吉, "HMM 音声合成を目的とした基本周波数パターン生成過程モデルのモデルパラメータ自動推定," *電子情報通信学会論文誌*, Vol.J98-D, No.3, pp.481-491(2015-3). 査読有  
DOI: 10.14923/transinfj.2014PDP0030
3. 越智景子, 広瀬啓吉, 峯松信明, "基本周波数パターン生成過程モデルの指令の差分に着目した発話の焦点制御," *電子情報通信学会論文誌*, Vol.J98-D, No.3, pp.524-533 (2015-3). 査読有  
DOI: 10.14923/transinfj.2014JDP7084
4. Oraphan Krityakien, Keikichi Hirose, and Nobuaki Minematsu, "Generation of fundamental frequency contours for Thai speech using the tone nucleus model," *Journal of Signal Processing, Research Institute of Signal Processing*, vol.16, no.4, pp.135-138 (2013-7). 査読有  
DOI: 10.2299/jsp.17.135
5. Qinghua Sun, Keikichi Hirose, and Nobuaki Minematsu, "A method for generation of Mandarin F0 contours based on tone nucleus model and superpositional model," *Speech Communication*, Vol.54, Issue 8, pp.932-945 (2012-10). 査読有  
DOI: 10.1016/j.specom.2012.03.005
6. Tatsuya Matsuda, Keikichi Hirose, and Nobuaki Minematsu, "Applying generation process model constraint to fundamental frequency contours generated by hidden-Markov-model-based speech synthesis," *Acoustical Science and Technology, Acoustical Society of Japan*, Vol.33, No.4, pp.221-228 (2012-7). 査読有  
DOI: 10.1250/ast.33.221

[学会発表](計34件)(下記に表記した18件以外に国内研究会での5件、国内全国大会での11件の発表がある。全て論文があるが、2以降の国際会議での発表はすべて査読有。いずれも筆頭著者が口頭発表を行った。)

1. 橋本哲弥, 柏木陽佑, 齋藤大輔, 広瀬啓吉, 峯松信明, "複数出力サブネットワークを有するディープニューラルネットワークに基づく声質変換," *情報処理学会研究報告, 音声言語情報処理研究会資料*, 2014-SLP-104(19)/電子情報通信学会音声研究会資料, SP2014-117, 東京工業大学すずかけ台キャンパス(神奈川県・横浜市), 2014年12月15-16日, pp.1-6.
2. Keikichi Hirose, Hiroya Hashimoto, Kyota Hyakutake, Daisuke Saito, and Nobuaki Minematsu, "Use of fundamental frequencies shaped by generation process model for HMM-based speech synthesis," *Proceedings IEEE International Conference on Signal Processing (ICSP'14)*, Hangzhou (China), October 19-23, 2014, pp.555-560. (Invited to Special Session)
3. Daisuke Saito, H. Doi, Nobuaki Minematsu, and Keikichi Hirose, "Voice conversion based on matrix variate gaussian mixture model," *Proceedings IEEE International Conference on Signal Processing (ICSP'14)*, Hangzhou (China), October 19-23, 2014, pp.567-576. (Invited to Special Session)
4. Daisuke Saito, Hidenobu Doi, Nobuaki Minematsu, and Keikichi Hirose, "Application of matrix variate Gaussian mixture model to statistical voice conversion," *Proceedings INTERSPEECH*, Singapore (Singapore), September 14-18, 2014, pp.2504-2508.
5. Keikichi Hirose, "Use of generation process model for controlling fundamental frequencies in HMM-based speech synthesis," *Proceedings of Forum Acusticum, Krakow (Poland)*, September 7-12, 6 pages in CD-ROM proceedings (ISSN 2221-3767). (Invited to Structured Session)
6. Daisuke Saito, Nobuaki Minematsu, and Keikichi Hirose, "Tensor representation for speaker characteristics in speech," *Proceedings of Forum Acusticum, Krakow (Poland)*, September 7-12, 5 pages in CD-ROM proceedings (ISSN 2221-3767). (Invited to Structured Session)
7. Ya Li, Jianhua Tao, Keikichi Hirose, Wei Lai, Xiaoying Xu, "Hierarchical stress generation with Fujisaki model in expressive speech synthesis," *Proceedings of International Conference on Speech Prosody*, Dublin (Ireland), May 20-23, 2014, pp.1032-1036.
8. Tomoyuki Mizukami, Hiroya Hashimoto, Keikichi Hirose, Daisuke Saito, and Nobuaki Minematsu, "Selection of training data for HMM-based speech synthesis from prosodic

- features - Use of generation process model of fundamental frequency contours -," Proceedings of International Conference on Speech Prosody, Dublin (Ireland), May 20-23, 2014, pp.1042-1046 .
9. Keikichi Hirose, "Control of fundamental frequencies in HMM-based speech synthesis using generation process model," Proceedings of International Symposium on Frontiers of Research on Speech and Music (FRSM-2014), Mysore (India), March 13-14, 2014, pp. 96-100. (Invited Speech)
  10. Hiroya Hashimoto, Keikichi Hirose and Nobuaki Minematsu, "Context labels based on "bunsetsu" for HMM-based speech synthesis of Japanese," Proceedings 8th ISCA Workshop on Speech Synthesis (SSW-8), Barcelona (Spain), August 31-September 3, 2013, pp.35-39.
  11. Oraphan Krityakien, Nobuaki Minematsu, and Keikichi Hirose, "Generation of fundamental frequency contours for Thai speech synthesis using tone nucleus model," Proceedings INTERSPEECH, Lyon (France), August 26-29, 2013, pp.1037-1041.
  12. Oraphan Krityakien, Keikichi Hirose, and Nobuaki Minematsu, "F0 contour generation of Thai speech using the tone nucleus model," Proceedings of RISP International Workshop on Nonlinear Circuits, Communication and Signal Processing (NCSP2013), Kona (U.S.A.), March 5-7, 2013, pp.81-84. (Student Paper Award)
  13. Daisuke Saito, Nobuaki Minematsu, Keikichi Hirose, "Tensor-based speaker space construction for arbitrary speaker conversion," Proceedings IEEE International Conference on Signal Processing (ICSP'12), Beijing (China), October 22-24, 2012, pp.595-598. (Invited to Special Session)
  14. Keikichi Hirose, Hiroya Hashimoto, Jun Ikeshima, and Nobuaki Minematsu, "Use of generation process model for synthesizing fundamental frequency contours in HMM-based speech synthesis," Proceedings IEEE International Conference on Signal Processing (ICSP'12), Beijing (China), October 22-24, 2012, pp.575-578. (Invited to Special Session)
  15. Hiroya Hashimoto, Keikichi Hirose, and Nobuaki Minematsu, "Improved automatic extraction of generation process model commands and its use for generating fundamental frequency contours for training HMM-based speech synthesis," Proceedings INTERSPEECH, Portland (U.S.A.), September 9-13, 2012, 4 pages in CD-ROM Proceedings.
  16. Daisuke Saito, Nobuaki Minematsu, Keikichi Hirose, "Effects of speaker adaptive training on tensor-based arbitrary speaker conversion," Proceedings INTERSPEECH, Portland (U.S.A.), September 9-13, 2012, 4 pages in CD-ROM Proceedings.
  17. Miaomiao Wang, Miaomiao Wen, Keikichi Hirose, and Nobuaki Minematsu, "Emotional voice conversion for mandarin using tone nucleus model – small corpus and high efficiency," Proceedings of International Conference on Speech Prosody, Shanghai (China), May 22-25, 2012, pp.163-166. Student paper award (2nd prize)
  18. Keikichi Hirose, Hiroya Hashimoto, Jun Ikeshima, and Nobuaki Minematsu, "Fundamental frequency contour reshaping in HMM-based speech synthesis and realization of prosodic focus using generation process model," Proceedings of International Conference on Speech Prosody, Shanghai (China), May 22-25, 2012, pp.171-174.
- 〔図書〕(計1件)
1. Keikichi Hirose & Jianhua Tao (eds.), Springer-Verlag, Berlin, "Speech Prosody in Speech Synthesis: Modeling and generation of prosody for high quality and flexible speech synthesis," (2015-4). 221 pages (Chapter 10: Use of generation process model for improved control of fundamental frequency contours in HMM-based speech synthesis, pp.145-160 執筆)
- 6 . 研究組織
- (1)研究代表者  
 広瀬 啓吉 (HIROSE, Keikichi)  
 東京大学・大学院情報理工学系研究科・教授  
 研究者番号 : 5 0 1 1 1 4 7 2
  - (2)研究分担者  
 峯松 信明 (MINEMATSU Nobuaki)  
 東京大学・大学院工学系研究科・教授  
 研究者番号 : 9 0 2 7 3 3 3 3
  - 齋藤 大輔 (SAITO Daisuke)  
 東京大学・大学院工学系研究科・助教  
 研究者番号 : 4 0 6 1 5 1 5 0