

別紙 HP 公開資料

戦略的国際科学技術協力推進事業（日本－インド研究交流）

1. 研究課題名：「音声言語における韻律のモデル化とそれによる音声合成」
2. 研究期間：平成 20 年 2 月～平成 24 年 3 月
3. 支援額： 総額 17,503,000 円
4. 主な参加研究者名：

日本側（研究代表者を含め 6 名までを記載）

	氏名	所属	役職
研究代表者	広瀬啓吉	東京大学	工学博士
研究者	藤崎博也	東京大学	工学博士
研究者	匂坂芳典	早稲田大学	工学博士
研究者	峯松信明	東京大学	博士(工学)
研究者	齋藤大輔	東京大学	科学修士
研究者	越智景子	東京大学	科学修士
参加研究者 のべ			13名

中国側（研究代表者を含め 6 名までを記載）

	氏名	所属	役職
研究代表者	DAS MANDAL, Shyamal Kr	C-DAC, Kolkata	Sr. Scientific Officer
研究者	DATTA, Asoke Kumar	C-DAC, Kolkata	教授
研究者	KUMAR, R. Ravindra	C-DAC, Kolkata	所長
研究者	SAHA, Arup	C-DAC, Kolkata	Scientific Officer
研究者	BASUE, Tulika	C-DAC, Kolkata	言語学者
研究者	WARSI, Anal	C-DAC, Kolkata	研究員
参加研究者 のべ			10名

5. 研究・交流の目的

本研究は、韻律のモデル化、音声合成の高度化、韻律モデルの実装等で、日印間の研究協力を深め、人間－機械間の音声コミュニケーション実現のための、両国間の協力関係を緊密にすることを目的とする。具体的には、インド側は、独自技術に基づくベンガル語の音声合成システムを開発しているが、抑揚の問題が浮き彫りになっている。これに対し、日本側が協力することにより、ベンガル語をはじめとするインド語の音声合成の高度化を達成する。特に、日印で協力して韻律分析を行い、韻律制御手法を確立して音声合成システムに実装する。本共同研究で日印が交流を通じて相互的に取り組むことで、将来に渡る音声技術での両国間の協力関係の構築が期待される。

6. 研究・交流の成果

6-1 研究の成果

日本側で開発し、その有効性を日本語を始めとする多くの言語について示している基本

周波数生成 (F0) パターン生成過程のモデル (F0 モデル) を用い、ベンガル語音声の韻律の分析を行った。その結果、ベンガル語音声では、フレーズ頭のアクセントが必ず負の値を取るといふ、他言語には見られない現象を確認した。分析結果を基に、与えられた文に対し、休止と F0 の時間変化パターン (F0 パターン) を生成する韻律制御を構築した。この韻律制御は、日本側で開発した日本語音声合成の韻律規則を基にしたもので、文の統語解析を行って得られる統語情報、語境界の位置情報、アクセントの種類などから、休止の生起、休止長、F0 モデルのパラメータを推定するものである。この韻律規則を、インド側で開発した音声合成システム ESNOLA に実装して、品質の高い合成音声を得た。

韻律制御の高度化の研究を日本側が中心となって進めた。HMM 音声合成は、柔軟な音声合成として、最近注目されているが、その韻律制御に問題があるとされる。これに対し、F0 モデルに基づく 2 つの改良を達成した。1 つ目は、学習音声コーパスの F0 パターンを F0 モデルで最良近似し、それを HMM の学習に用いるもので、2 つ目は HMM 音声合成で生成される F0 パターンを F0 モデルの枠組みで修正するものである。合成音声の聴取により、音質向上を確認した。なお、F0 パターンを F0 モデルで表現することで、談話の焦点、意図・態度などとの明確な対応が得られ、韻律の柔軟な制御が可能となる。実際に、F0 モデルパラメータの差分に着目した談話の焦点の実現を音声合成で達成した。

なお、本研究交流の広報の場として、2009 年 11 月に International Workshop on Spoken Language Prosody をインド側 C-DAC で開催した。また、2010 年 9 月に幕張メッセで開催された INTERSPEECH 2010 でデモブースを設け、開発したベンガル語音声合成のデモなど、研究成果を広報した。

6-2 人的交流の成果

インド側研究代表者の Das Mandal 氏、研究メンバーの Saha 氏、Warsi 氏は、いずれも若手研究者であり、研究交流だけでなく、研究の進め方、論文の書き方なども指導した。特に、Das Mandal 氏は大きく飛躍し、インドの工学分野の有力校である Indian Institute of Technology Kharagpur 校の助教授に迎えられた。その後も、音声合成に留まらず、音声認識、発音教育などをテーマとして、緊密な研究協力関係を保っている。

本研究交流メンバー以外のインド人研究者との交流も深めることが出来た。特に KIIT Group of Colleges の Agrawal 教授とは、ヒンズー語など、他方言への研究の展開を継続的に討議している。

7. 主な論文発表・特許等 (5 件以内)

相手国側との共著論文については、その旨を備考欄にご記載ください。

論文 or 特許	・論文の場合： 著者名、タイトル、掲載誌名、巻、号、ページ、発行年 ・特許の場合： 知的財産権の種類、発明等の名称、出願国、出願日、 出願番号、出願人、発明者等	備考
論文	T. Matsuda, K. Hirose, N. Minematsu, "HMM-based F0 contour synthesis using the generation process model," Acoustical Science and Technology, Acoustical Society of Japan, 採録決定 (2012).	
論文	M. Wang, K. Hirose, and N. Minematsu, "Improvement of prosody in HMM-based speech synthesis using generation process model," Journal of Research Institute of Signal Processing, vol.15, no.4, pp.279-282 (2011-7).	
論文	K. Hirose, "Fundamental frequency contour generation process model for improved and flexible control of prosodic features in hmm-based speech synthesis," Proceedings of International Symposium on Frontiers of Research on Speech and Music (FRSM-2012), Gurgaon, pp.302-306 (2012-1).	招待論文
論文	A. Warsi, T. Basu, K. Hirose and H. Fujisaki, "Prosodic Comparison of	共著

	Declarative and Interrogative Utterances in Standard Colloquial Bangla,” Proceedings Oriental COCOSDA, Hsinchu, 6 pages in CD-ROM Proceedings (2011-10).	
論文	S. Das Mandal, A. Saha, T. Basu, H. Fujisaki, and K. Hirose, “ Modeling of intra-sentence pauses in Bangla readout speech: Occurrences and duration,” INTERSPEECH 2010, Makuhari, pp.1764-1767 (2010-9).	共著