

大学名	東京大学		
University	The University of Tokyo		
外国人研究者	モッラ カデマル イスラム		
Foreign Researcher	MOLLA, Md. Khademul Islam		
受入研究者	広瀬啓吉	職名	教授
Research Advisor	HIROSE, Keikichi	Position	Professor
受入学部/研究科	情報理工学系研究科		
Faculty/Department	Graduate School of Information Science and Technology		

<外国人研究者プロフィール/Profile>

国籍	バングラデシュ
Nationality	Bangladesh
所属機関	ラジャヒ大学
Affiliation	The University of Rajshahi
現在の職名	教授
Position	Professor
研究期間	2015年1月24日～2015年3月26日
Period of Stay	January 24, 2015 - March 26, 2015
専攻分野	音響・音声信号処理
Major Field	Acoustic and speech signal processing



外国人研究者による講演の様子

<外国人研究者からの報告/Foreign Researcher Report>

①研究課題 / Theme of Research
Speech Enhancement for Advanced Man-Machine Interaction
②研究概要 / Outline of Research
In soft-thresholding, speech enhancement scheme, speech is decomposed into sub-bands, and their threshold factors are estimated from the noise variance. In the current work, bivariate empirical mode decomposition (BEMD) is introduced to decompose noisy speech and fractional Gaussian noise (fGn) together into a set of sub-bands called intrinsic mode functions (IMFs).
③研究成果 / Results of Research
The effectiveness of the proposed BEMD-based algorithm is tested using computer simulation with different male and female utterances (English sentences) randomly selected from TIMIT Database. The white noise is added to the clean speech to obtain the noisy speech signals at different noise levels. From the result, it is clear that the method has advantages over conventional methods.
④今後の計画 / Further Research Plan
Further research is planned to apply the enhanced speech in the improvement of speech recognition performance. Denoising in real environment will be conducted in this circumstance. Also, efficient estimation of fundamental frequency and voice activity detection are the scope of future research. These may contribute a lot in improving speech synthesis and recognition performances.

<受入研究者からの報告/Research Advisor Report>

①研究課題 / Theme of Research

音声強調によるマン・マシン・インターフェースの高度化

②研究概要 / Outline of Research

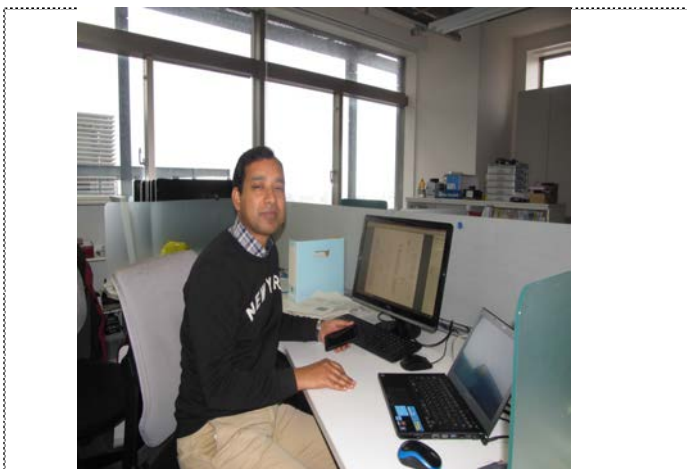
当初は音声の位相情報も考慮した音源分離を中心に研究を進める予定であったが、より実効のあるsoft-thresholdingに焦点を当てるように指導した。帯域分割を行って、それぞれの帯域毎に適切な域値を雑音の分散から求めるものであるが、本研究では、帯域分割と域値の設定にbivariate empirical mode decomposition (BEMD)を利用して性能向上を図ることとした。BEMDは雑音音声と雑音を同時に処理し、それぞれを、intrinsic mode functionに分解するもので、分解された成分同士の対応が保障される。分解された成分は、ある時点について見れば、正弦波となっており、信号の特徴に即した適応的な帯域分割が実現される。分解された雑音から域値の適切な設定が可能となり、soft-thresholdingの性能向上が期待される。さらに、この音声強調により、音声認識の性能向上を試みる。

③研究成果 / Results of Research

提案手法の有効性を、TIMITデータベースの男性と女性話者音声について実験的に調べた。雑音のない音声に、白色雑音を種々のSNRIになるように付加して雑音音声とし、実験に用いた。スペクトログラムおよびSNR低減率を評価指標として、既存のsoft-thresholding手法との比較を行った。その結果、0dBから30dBのSNRの範囲で、提案手法の有効性が確認された。soft-thresholdingでは、音声信号を歪ませる危険性があるが、perceptual evaluation of speech quality (PESQ)を指標とした主観評価実験により、提案手法は、この観点からも従来手法と比べ優れていることが明らかとなった。音声強調された音声を音声認識に用いることで、その性能向上が得られる見通しを得た。なお、滞在中にEMDに関する研究紹介を受け入れ研究者研究室の学生に対して行った。

④今後の計画 / Further Research Plan

音声強調された音声をを用いた音声認識の実験を進める。認識の前に、SPLICEなどのパラメータレベルの音声強調も行うことを想定し、音声認識に最適なsoft-thresholdingの域値の設定について詳細な検討を行う。さらに、白色雑音を用いた計算機シミュレーションに止まらず、実環境での雑音音声を対象とした実験を行う。帰国後も、電子メール等によって緊密な交流を保ちながら研究を進める。6月には、外国人研究者のRajshahi大学で行われる国際会議に出席する予定であり、研究の方向について詳細に討議する。なお、滞在中の成果として、国際会議INTERSPEECHに論文を投稿した。



研究室での外国人研究者



ポーランドからの研究者も交えて懇親