

研究種目：基盤研究（A）

研究期間：2007～2010

課題番号：19200017

研究課題名（和文） 聴覚・音声機能の支援・拡張技術に関する総合的研究

研究課題名（英文） Comprehensive research on extension and supporting technologies for auditory and speech functions

研究代表者 河原 英紀（KAWAHARA HIDEKI）

和歌山大学・システム工学部・教授

研究者番号：40294300

研究代表者の専門分野：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・ユニバーサルデザイン・QOL・能力の拡張

1. 研究計画の概要

本研究では、衰えた聴力や声を取り戻すだけでなくそれらをより優れたものとするための支援技術の基盤確立を目的とする。本研究の狙いは、衰えた能力の機械による貧弱な模倣と置き換えにではなく、それぞれの個人が有する能力を自然に拡張する形で支援するところにある。

2. 研究の進捗状況

本研究の基盤技術である高度な音声分析変換合成術 STRAIGHT が、初年度の研究で生まれた発明により、TANDEM-STRAIGHT として、本質的な進歩を遂げた技術として生まれ変わったことが、本研究の順調な進捗に大きく貢献している。また、その基盤の上に構築される応用技術である音声変換手法も大きく進展した。それらの大きな進歩により生まれた、当初の目標をより高いレベルで実現できる可能性をも追求することで、本研究は、当初計画以上に順調に進捗していると評価できる状況にある。

以下、当初計画で挙げた、それぞれの研究項目に対応させて補足説明する。

(1) 個人性、自然性を高い品質で保存した、音声の明瞭度、声年齢等諸属性の改善処理技術については、前節で説明した発明を順調に取り込んだ技術開発が進んでおり、当初計画

(2) 実環境下での複数点受音による、主観的品質と明瞭度を改善できる雑音抑圧／選択的聴取技術

(3) 記憶された音声と同じ印象を与える音声の記述及び変換合成技術

3. 現在までの達成度

前節で説明した様々な要素を総合的に判断すると、本研究の現在までの達成度の区分

を、以下のように評価することができる。

①当初の計画以上に進展している。

この評価には、繰返しになるが、本研究の基盤技術である STRAIGHT が、新発明により、TANDEM-STRAIGHT という、理論的にも実用上も優れた技術として生まれ変わったことが大きく貢献している。

4. 今後の研究の推進方策

今後は、当初の計画で提案したように、成果を社会に還元できるレベルにまで整備することを目指すとともに、様々な応用イメージを具体的に提示することを重点として進める。

5. 代表的な研究成果

（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 13 件）

①Laetitia Bruckert, Patricia Bestelmeyer, Marianne Latinus, Julien Rouger, Ian Charest, Guillaume A. Rousselet, Hideki Kawahara, Pascal Belin, Vocal Attractiveness Increases by Averaging, Current Biology, Volume 20, Issue 2, 116-120, 26 January 2010.

②森勢 将雅, 高橋 徹, 河原 英紀, 入野 俊夫, “分析時刻に依存しない周期信号のパワースペクトル推定法を用いた音声分析,” 電子情報通信学会論文誌, Vol. J92-A, No. 3, pp. 163-171, Mar. 2009.

③ 河原英紀: 音声モーフィングの背景と可能性、音声言語医学、50(2), pp. 131-135, 2009.

④ 森勢 将雅, 高橋 徹, 河原 英紀, 入野 俊夫, “窓関数による分析時刻の影響を受けに

くい周期信号のパワースペクトル推定法”, 電子情報通信学会論文誌, Vol. J90-D, No. 12, pp. 3265-3267, 2007.

〔学会発表〕 (計 106 件)

- ① H. Doi, K. Nakamura, T. Toda, H. Saruwatari, K. Shikano. Statistical approach to enhancing esophageal speech based on Gaussian mixture models. Proc. ICASSP, pp. 4250-4253, Dallas, USA, Mar. 2010.
- ② Ryoji Okamoto, Yu Takahashi, Hiroshi Saruwatari, Kiyohiro Shikano, MMSE STSA Estimator with nonstationary noise estimation based on ICA for high-quality speech enhancement, Proc. ICASSP 2010, pp. 4778-4781, March 2010.
- ③ Hideki Kawahara, Ryuichi Nisimura, Toshio Irino, Masanori Morise, Toru Takahashi and Hideki Banno: High-quality and light-weight voice transformation enabling extrapolation without perceptual and objective breakdown, ICASSP2010, 4818-4821 (March, 2010)
- ④ Hideki Kawahara, Toru Takahashi, Masanori Morise and Hideki Banno: Development of exploratory research tools based on TANDEM-STRAIGHT, Proc. APSIPA, Sapporo, pp. 111-120 (2009).
- ⑤ K. Nakamura, T. Toda, H. Saruwatari, K. Shikano. Electrolaryngeal speech enhancement based on statistical voice conversion. Proc. INTERSPEECH, pp. 1431-1434, Brighton, UK, Sep. 2009.
- ⑥ Hideki Kawahara, Masanori Morise, Toru Takahashi, Hideki Banno, Ryuichi Nisimura and Toshio Irino: Observation of empirical cumulative distribution of vowel spectral distances and its application to vowel based voice conversion, Proc. Interspeech2009, Brighton, pp. 2647-2650 (2009).
- ⑦ H. Kawahara, R. Nisimura, T. Irino, M. Morise, T. Takahashi, B. Banno, Temporally variable multi-aspect auditory morphing enabling extrapolation without objective and perceptual breakdown, Proc. ICASSP, 3905-3908, 19-24 (2009).
- ⑧ Yuji Kubota, Kazunori Komatani, Tetsuya Ogata, Hiroshi G. Okuno: Design and Implementation of 3D Auditory Scene Visualizer towards Auditory Awareness with Face Tracking, Proceedings of IEEE International Symposium on Multimedia (ISM08), pp. 468-476, Berkeley, Dec. 16. 2008.
- ⑨ Ryota Takafuji, Yoshimitsu Mori,

Hiroshi Saruwatari, Kiyohiro Shikano, Binaural Hearing-Aid System Using SIMO-Model-Based ICA and Directivity-Dependency-Reduced Binary Masking, International Conference on Signal Processing (ICSP'08), Beijing, China, pp. 320-323, Oct. 2008.

⑩ Keigo Nakamura, Tomoki Toda, Yoshitaka Nakajima, Hiroshi Saruwatari and Kiyohiro Shikano, Evaluation of Speaking-Aid System with Voice Conversion for Laryngectomees Toward Its Use in Practical Environments, INTERSPEECH, pp. 2209-2212, Brisbane, Australia, Sep. 2008.

⑪ Hideki Kawahara, Masanori Morise, Toru Takahashi, Ryuichi Nisimura, Toshio Irino, Hideki Banno, TANDEM-STRAIGHT: A temporally stable power spectral representation for periodic signals and applications to interference-free spectrum, F0, and aperiodicity estimation, Proc. ICASSP 2008, Las Vegas, pp. 3933-3936 (2008)

⑫ Hideki Kawahara, Remaking Speech Revisited -- STRAIGHT and TANDEM-STRAIGHT and Their Implications, Asian Workshop on Speech Science and Technology, Tokyo, Japan, March 20, 2008

〔図書〕 (計 5 件)

⑬ Kazuhiko Kakehi, Yuko Sogabe and Hideki Kawahara, "Research on emotional perception of voices based on a morphing method," Ed. Krzysztof Izdebski, Emotions in the human voice volume III, Plural Publishing, pp. 1-14, 2008.

〔産業財産権〕

○出願状況 (計 0 件)

〔その他〕

受賞: 計 6 件