Abstract

This paper proposes a voice morphing system for people suffering from Laryngectomy, which is the surgical removal of all or part of the larynx or the voice box, particularly performed in cases of laryngeal cancer. A primitive method of achieving voice morphing is by extracting the source's vocal coefficients and then converting them into the target speaker's vocal parameters. In this paper, we deploy Gaussian Mixture Models (GMM) for mapping the coefficients from source to destination. However, the use of the traditional/conventional GMM-based mapping approach results in the problem of over-smoothening of the converted voice. Thus, we hereby propose a unique method to perform efficient voice morphing and conversion based on GMM, which overcomes the traditional-method effects of over-smoothening. It uses a technique of glottal waveform separation and prediction of excitations and hence the result shows that not only over-smoothening is eliminated but also the transformed vocal tract parameters match with the target. Moreover, the synthesized speech thus obtained is found to be of a sufficiently high quality. Thus, voice morphing based on a unique GMM approach has been proposed and also critically evaluated based on various subjective and objective evaluation parameters. Further, an application of voice morphing for
Laryngectomees which deploys this unique approach has been recommended by this paper.

References

- Abe M., Nakamura S., Shikano K. and Kuwabara H., "Voice conversion through vector quantization" International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 1988, 655-658.
- Baudoin G. and Stylianou Y., "On the transformation of the speech spectrum for voice conversion" International Conference on Spoken Language (ICSLP), Philadelphia, October 1996, Vol. 3, 1405-1408.
- Kain A. and Macon M., "Spectral voice conversion for text to speech synthesis" Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 1998, Vol. 1, 285-288.
- Stylianou Y. and Cappe O., "A system for voice conversion based on probabilistic classification and a harmonic plus noise model" International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 1998, Seattle, 281-284.
- Ye H. and Young S., "High quality voice morphing" International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2004, Montreal, Vol. 1, 9-12.
- Upperman, G., "Linear Predictive Coding In Voice Conversion", December 21, 2004.
- Bradbury J., "Linear Predictive Coding" December 5, 2000.
- Gundersen, T., "Voice Transformation based on Gaussian mixture models" Master of Science in Communication Technology Thesis, Norwegian University of Science and Technology, Department of Electronics and Telecommunications, 2010, 55.
- Cliff M., "GMM and MINZ Program Libraries for Matlab" Krannert Graduate School of Management, Purdue University, March 2, 2003.
- Scherrer B., "Gaussian Mixture Model Classifiers" February 5, 2007.
- Resch B., "Mixtures of Gaussians-A Tutorial for the Course Computational Intelligence" Signal Processing and Speech Communication Laboratory Inffeldgasse 16c, http://www.igi.tugraz.at/lehre/CI, June 2012.
- XuN. and Yang Z., "A Precise Estimation of Vocal Tract Parameters for High Quality Voice Morphing" 9th International Conference on Signal Processing (ICSP), October 2008, 684-687.
- Nakamura K., Toda T., Nakajima Y., Saruwatari H. and Shikano K., "Evaluation of Speaking-Aid System with Voice Conversion for Laryngectomees Toward Its Use in Practical Environments" Interspeech (ISCA), 2008, Brisbane.
- Huang X., Acero A. and Hon H. "Spoken Language Processing: A Guide to Theory, Algorithm and System Development" Prentice Hall, 2001.
- Reynolds D., "Gaussian Mixture Models" Encyclopedia of Biometrics, 2009, 659-663.
- Mesbahi L., Barreaud V, and Boeffard O., "GMM-Based Speech Transformation Systems under Data Reduction" Sixth ISCA Workshop on Speech Synthesis, 2007, 119-124.
Analysis of a Modern Voice Morphing Approach using Gaussian Mixture Models for Laryngectomees

- CMU_ARCTIC Speech Synthesis Databases, Carnegie Mellon University, http://festvox.org/cmu_arctic, March 2012.
- Russell M. , "Towards Speech Recognition using Palato-Lingual Contact Patterns for Voice Restoration," PhD Thesis, Faculty of Engineering, University of the Witwatersrand, June 2011.
- Dae-Hyeong K. et al. , "Epidermal Electronics," Science, Vol. 333, No. 6044, 12 August 2011, 838-843.

Index Terms

Computer Science          Signal Processing

Keywords

Voice Morphing  Laryngectomy  Gaussian Mixture Models