

O-COCOSDA 2010 Philippine Country Report

Jocelynn Cu¹, Joel Ilaio^{1,2} and Ethel Ong¹

¹College of Computer Studies, De La Salle University

²Electrical and Electronics Institute, UP-Diliman

{jiji.cu;joel.ilao;ethel.ong}@delasalle.ph

1. Country Report

Recent works on Filipino Speech Synthesis investigated the effect of incorporating intensity, duration and pitch modification models, on the performance of Filipino speech synthesizers [1]-[2]. ASRs for Filipino speech are applying NLP techniques in order to improve their performance [3]. For this purpose, a large Filipino text database was built by harvesting documents from World Wide Web using a Query-based scheme in order to improve statistical language models for Filipino [4]. Filipino ASR system based on the CMU Sphinx Framework was used to perform closed captioning of Filipino Video News Broadcasts [5]. Empirical studies were also done on Filipino speech rhythm [6], and feature space reduction for faster speech recognition [7]. Some completed works focused on real-time implementation of speech technology-aligned techniques using DSP-specialized hardware [8]-[10], development of novel algorithms for speech coding and voice enhancement in Next Generation Networks (NGN) [11]-[14], a vowel classification system of Filipino Sung speech for music education [15], and tools for automatic laughter segmentation and affect classification [16]-[17]. A Filipino Multimodal Emotion Database [18] containing acted and simultaneous Filipino emotional speech was also created to support research on affective systems. The bi-directional hybrid machine translation system for English and Filipino is now made available for public use through the Language Grid Playground [19]. On-going projects include automatic story generation and the eLegislation and eParticipation system [20] for the Philippine Senate Blue Ribbon Committee. The latter, funded by PANeGOV, is a system prototype created for the BRC stakeholders (e.g. general public, legislative staff) to extract, retrieve and organize existing and future data for easier reference and use.

2. References

- [1] L.R. Lazaro and L. Policarpio. "Incorporating Duration and Intonation Models in Filipino Speech Synthesis," Undergraduate Thesis, UP Diliman, 2009.
- [2] J.D. Albete, N.M.V. David, L.R.S. Lazaro, and R.C.L. Guevara, "Enhancements in Prosody and Smoothness of Synthesized Filipino Speech", Proc. of ISMAC 2010, Manila, Phils., Sept. 8-9, 2010.
- [3] R.M. Almonte, M.G.A.R. Bayona, P.S.T. Corrales, "Incorporating natural language processing techniques in Filipino speech recognition", Undergraduate Thesis, 2009.
- [4] J.P. Ilaio and R.C.L. Guevara, "Mining Filipino-English Corpora from the Web", International Symposium on Multimedia and Communication Technology, Manila, Sept. 8-9, 2010.
- [5] M.C.V. Burgos and M.S. De Lara. "Closed Captioning System in Filipino", U.P. Diliman, Undergraduate Thesis, 2010.
- [6] T. Santos, "Filipino Speech Rhythm Classification Based on Computational and Perceptual Features", U.P. Diliman, Undergraduate Thesis, 2010.
- [7] J.P. Ilaio and R.C.L. Guevara, "Filipino Speech Phoneme Classification Using a Reduced Feature Set for Multi-Layer Perceptron and Support Vector Machine Classifiers", 5th ERDT Conference, Sept. 10, 2010.
- [8] A.D.D. Go, "Implementation of a Variable Step Size NLMS Algorithm for Speech Noise Reduction on an ADSP-21065L EZ Kit Lite Board", UP Diliman, Undergraduate Thesis, 2008.
- [9] A.U.Q. Unisa and R.C.L. Guevara, "Real-time implementation of wideband sinusoidal speech coder on ADSP-21065L", Santorini, Greece, 2009.
- [10] K.K.A. Medina, "Implementation of a Least Mean Square (LMS) Based ITU-T G.168 Digital Network Echo Canceller on Xilinx Virtex 5 SXT FPGA", UP Diliman, Undergraduate Thesis, 2009.
- [11] F.M. Ang and R.C.L. Guevara, "Joint Source-Channel Coding for Packet Network Transmission of Low Bit-Rate Encoded Wideband Speech", 3rd ERDT Conference, Manila, Philippines, 2009.
- [12] F.M. Ang and R.C.L. Guevara, "A robust packet loss recovery scheme for wideband speech codecs", Proc. of 9th Int'l Conf. on Communications and Information Technologies, Incheon Korea, pp. 531-536, 2009.
- [13] M.O. Cordel and R.C.L. Guevara, "Parameter-Based Voice Enhancement for Mobile Communication", Proc. of ISMAC 2010, Manila, Sept. 8-9, 2010.
- [14] M.O. Cordel and R.C.L. Guevara, "A Parameter-Based Echo Canceller for Wideband Adaptive Multirate Codec", 5th ERDT Conference, Sept. 10, 2010.
- [15] V.B. Bustos, T.J.G. De la Cruz, R.M.G. Acoymo, R.C.L. Guevara, "Development of Feature Set, Classification Implementation and Applications for Vowel Migration/Modification in Sung Filipino (Tagalog) Texts and Perceived Intelligibility", Science Diliman, 21:2 (2009).
- [16] H. P. Bantiling, S. R. Gadi, J. C. Lee, J. V. Yang, and M. T. Suarez, "Automatic Video Segmentation Tool for Laughter Detection Based on Audio Features", 15th OU-DLSU Joint Academic Research Workshop, Sept. 29 – 30, 2010.
- [17] J. Alonzo, J. Campita, S. Lucila, M. Miranda, and M. T. Suarez, "Discovering Emotions in Filipino Laughter", 15th OU-DLSU Joint Academic Research Workshop, Sept. 29 – 30, 2010.
- [18] J. Cu, M. T. Suarez, and M. Sta. Maria, "A Filipino multimodal emotional database," Proc. LREC Int'l Workshop on Multimodal Corpora 2010, Malta, May 2010.
- [19] <http://langrid.org/playground/translation.html>
- [20] <http://panegov.net/projects/project1.htm>