FEDERICO ANG

Speech / Multimedia Signal Processing & Voice AI

fed.dsp@gmail.com | +81 90 3699 3536 | Tokyo, Japan (PR) | fmang | feddybear | angfederi.co

TECH STACK

Python | Shell | Open Source | Matlab | C++ | Cloud | PyTorch | Docker | k8s | git | LaTeX | APIs

TECH SKILLS

Signals and systems | Speech processing | Voice AI | IVR systems | Audio manipulation | NLP/NLU | GPT/LLMs | Machine Learning | Data Structures | Algorithms | Probability | Circuit Theory | Communications | Networking | Cloud | Scalability | MLOps | Security and privacy | Productionization

LANGUAGES

Filipino: Native
English: Fluent
Japanese: Mid-level

REFERENCES

Available upon request.

ABOUT ME

Dedicated R&D scientist in speech signal processing, specializing in AI-driven voice and conversational systems. Contributed to research in speech recognition and synthesis, and in developing practical voice AI systems for consumer applications.

EXPERIENCE

Assistant Manager | Rakuten Group, Inc.

Jan 2020 – Jun 2024 | Tokyo, Japan

- Performed managerial, R&D (including scholastic), and engineering roles surrounding speech and conversational-AI-related tasks in the company.
- Conducted hiring interviews, on-boarded and mentored new members and interns, published and presented for a conference paper on speech synthesis, delivered presentations at both departmental and company-wide events.
- Led a project to improve a family-safe spoken-word filter for our online selling platform, conducted comprehensive research and survey to identify the optimal dataset to train and boost a custom in-house speech recognition model for flagging inappropriate words. This could have cut costs significantly by reducing or entirely removing manual video checks. The project's progression to production was constrained by budgetary limits, but it had significant potential for handling more sellers efficiently.
- Pioneered the 'Speech Factory' initiative, advocating for changes to the development of speech recognition models within the company. My approach involved creating a holistic development platform integrating MLOps practices for all steps in the pipeline: semi-automated annotation, use-case-driven customization, and leaderboard-style benchmarking system. This was tested on an actual IVR use case in production (Nuance-based) that resulted in a 7x increase in annotation and optimization efficiency and significant cost savings, while providing a scalable ASR model development environment.

Research Scientist | Rakuten Group, Inc.

Jan 2018 – Jun 2024 | Tokyo, Japan

- Worked on baseline training and finetuning of several modern speech AI frameworks (NeMo, ESPnet, icefall, speechbrain) and querying APIs (cloud-based and AI products) for maintaining an internal running Japanese ASR performance scoreboard and develop PoCs related to speech tech: translation with ASR and TTS conversion.
- Improved the speech recognition accuracy of an internal HMM-DNN ASR model for voice-based product search by ~20% absolute, and a proprietary model (Nuance) for an IVR use case by ~7% absolute.
- Developed the voice interface of a concierge robot for a conference demonstration interacted by ~50 people.

Assistant Professor | University of the Philippines Diliman

Jan 2010 – Mar 2013, Aug 2016 – Jul 2017 | Quezon City, Philippines

- Served various roles from teaching, research advising, lab directing, and organizing university-wide events.
- Taught university-level engineering courses from circuit theory to advanced signal processing, paneled for and advised successful undergraduate student projects:
  - Sebastian, A.R., "Improving Filipino Speech Synthesis Using Voice Conversion," 2017.
  - Burgos, M.C. and M. De Lara, "Automatic Speech Recognizer for Closed Captioning of Filipino News Broadcasts," 2010.
  - Eala, M.C.F., "A Prototype for Hands-Free Filipino Text Messaging," 2011.
  - Ancheta, J.C.M., K. Chua, and K.M. Francia, "An Evaluation of Smoothing Techniques for Language Modeling in Filipino Automatic Speech Recognition Systems," 2012.
  - Crisostomo, R.L. and L. Godoy, "Improvements on the Development of a Closed Captioning System for Filipino News Broadcasts," 2012.

Other Posts

- **Temporary post as Assistant Professor (Apr 2016 – Aug 2016)** at the College of Computer Studies, **De La Salle University, Manila, Philippines**. Taught courses in computer architecture, feedback control systems and web security.
- **Guest researcher (Jul 2009 – Dec 2009)** at the International Center for Advanced Communication Technologies (interACT) in **Karlsruhe Institute of Technology, Karlsruhe, Germany**. Underwent training related to large-scale automatic speech recognition development.

EDUCATION

PhD in Media and Network Technologies | Hokkaido University

Apr 2013 – Mar 2016 Sapporo, Japan

- Coursework on data mining and retrieval, AI, pattern recognition, statistical learning, and multiagent systems.
- Research assistant at the Information and Communication Networks (ICN) Laboratory
- Dissertation entitled "A Study of Time-Varying Speech Features in the Context of Noise-Robust Speech Recognition" under the supervision of Dr.Eng. Yoshikazu Miyanaga.

MS in Electrical and Electronics Engineering | University of the Philippines Diliman

Nov 2007 – Apr 2009 Quezon City, Philippines

- Coursework on linear systems, advanced digital and wireless communication systems, advanced digital and adaptive signal processing, advanced networking, coding and information theory, machine learning, and natural language processing.
- Contributed to several funded research projects as an affiliate university research assistant for the Digital Signal Processing (DSP) Laboratory
- Master's thesis entitled "Joint Source-Channel Coding for Packet Network Transmission of Low Bit-Rate Encoded Wideband Speech" under the supervision of Dr. Rowena Cristina L. Guevara.

BS in Computer Engineering | University of the Philippines Diliman

Jun 2003 – Nov 2007 Quezon City, Philippines

- Digital signal processing laboratory trainee and affiliate
- Undergraduate student project entitled "On-Device (Symbian S60) Implementation of an Automatic Filipino Speech Recognition System"

Massive Online Open Courseware Certificates

- Structuring Machine Learning Projects, Coursera, Oct 2017.
- Improving Deep Neural Networks: Hyperparameter Tuning, Regularization, and Optimization, Coursera, Sep 2017.
- Neural Networks and Deep Learning, Coursera, Sep 2017.
- Discrete-Time Signal Processing, Graduate-Level DSP Group, MIT 6341x, EdX, May 2015.
- Artificial Intelligence, UC Berkeley, EdX, May 2015.
- Discrete-Time Signals and Systems, Part 1: Time-Domain, Rice Uni, EdX, Mar 2015.
- Introduction to Artificial Intelligence, Stanford Uni, Udacity, Dec 2011.

CONFERENCES

Conference Speaker

- 22nd Conference of the **Oriental COCOSDA (International Committee for the Coordination and Standardisation of Speech Databases and Assessment), Cebu, Philippines, Oct 2019**
- **High Impact Technology Solutions (HITS) Forum**, 2011 National Science and Technology Week Celebration, **Pasay City, Philippines, Jul 2011**.

Conference Session Chair

- 22nd Conference of the **Oriental COCOSDA (International Committee for the Coordination and Standardisation of Speech Databases and Assessment), Cebu, Philippines, Oct 2019**
- **IEEE Region 10 Conference (IEEE TENCON), Cebu City, Philippines, Nov 2012**
- 7th **International Conference on Natural Language Processing and Knowledge Engineering (NLP-KE), Tokushima, Japan, Nov 2011**.

Research Presenter

- **IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Online, Jun 2023**.
- 20th **International Conference on Digital Signal Processing (DSP), Singapore, Jul 2015**.
- **International Symposium on Circuits and Systems (ISCAS), Melbourne, Australia, Jun 2014**.
- **IEEE Region 10 Conference (IEEE TENCON), Cebu City, Philippines, Nov 2012**.
- 7th **International Conference on Natural Language Processing and Knowledge Engineering (NLP-KE), Tokushima, Japan, Nov 2011**.
- 4th **AUN/SEED-Net Regional Conference on Information and Communication Technology (RCICT), Ho Chi Minh City, Vietnam, Oct 2011**.
- 9th **International Symposium on Communications and Information Technologies (ISCIT), Incheon, South Korea, Sep 2009**.
- 4th **International Colloquium on Signal Processing and its Applications (CSPA), Kuala Lumpur, Malaysia, Mar 2008**.

PUBLICATIONS

- Xin, Detai, S. Adavanne, F. Ang, A. Kulkarni, S. Takamichi, and H. Saruwatari, **Improving Speech Prosody of Audiobook Text-to-Speech Synthesis with Acoustic and Textual Contexts**. In: IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Rhode Island, Greece, 2023.
- Ang, F.M., H. Tsutsui, and Y. Miyanaga, **Incorporation of Time-Varying LP Cepstral Features in HMM-Based Isolated Word Speech Recognition**. In: International Symposium on Signals, Circuits and Systems (ISSCS), Iaşi, Romania, 2015.
- Ang, F.M., Y. Miyanaga, R.C.L. Guevara, R. Cajote, M.G.A. Bayona. **Open Domain Continuous Filipino Speech Recognition with Code-Switching**. In: 2014 International Symposium on Circuits and Systems (ISCAS), Melbourne, Australia, 2014.
- Ang, F.M., J.C.M. Ancheta, K.M. Francia, and K. Chua. **Evaluation of Smoothing Techniques for Language Modeling in Automatic Filipino Speech Recognition**. In: The 2012 IEEE Region 10 Conference, Cebu City, Philippines, 2012.
- Ang, F.M., M.C. Burgos, and M. De Lara. **Automatic Speech Recognition for Closed-Captioning of Filipino News Broadcasts**. In: 7th International Conference on Natural Language Processing and Knowledge Engineering (NLP-KE), Tokushima, Japan, 2011.
- Ang, F.M., R. Almonte, M.G.A. Bayona, and L.R. Lazaro. **A Summary of Past and Current Developments in Filipino Speech Recognition and Speech-to-Text**. In: Proceedings of the 4th AUN/SEED-Net Regional Conference on Information and Communication Technology (RCICT), Ho Chi Minh City, Vietnam, 2011.
- Ang, F.M., and R.C.L. Guevara. **A Robust Packet Loss Recovery Scheme for Wideband Speech Codecs**. In: Proceedings of the 9th International Symposium on Communications and Information Technologies (ISCIT), Incheon, South Korea, 2009.
- Ang, F.M., and R.C.L. Guevara. **On-Device Implementation of an Automatic Filipino Speech Recognition System**. In: Proceedings of the 4th International Colloquium on Signal Processing and its Applications (CSPA), Kuala Lumpur, Malaysia, 2008.

PROJECTS (PUBLIC ACCESS)

FlipVox: Filipino ASR Suite (on-going) |

2024

- FlipScribe: Filipino Video Subtitling Tool
- FlipGraph: Streaming Filipino Dictation Tool
- FlipStar: Filipino Voice AI Leaderboard

Japanese Audiobook TTS Project Demo |

2022

- Combined bilateral lexical and audio context embeddings for better modeling of prosody in longform, multi-speaker TTS

Filipino (Tagalog) ASR model recipe for kaldi (featured in Alpha Cephei's vosk) | |

2020

- Data augmentation via noise, volume and tempo perturbation using MUSAN dataset.
- Language modeling via SRILM