Hieu-Thi Luong

E-mail : contact (at) hieuthi.com
Github : https://github.com/hieuthi
Linkedin : https://www.linkedin.com/in/luonghieuthi/
Website : https://www.hieuthi.com/

Experience

Project Researcher

National Institute of Informatics - Yamagishi Lab | Tokyo, Japan

October 2020 - present

Focus on Voice Cloning of Text-to-Speech and Voice Conversion for the "Voice Personae: Speaker Identity Cloning And Protection". Develop and expand speech synthesis further to capture personality, non-vocalization expression, etc.

Research Assistant

National Institute of Informatics - Yamagishi Lab | Tokyo, Japan

November 2017 - August 2020

Statistical parametric speech synthesis systems, multi-speaker models, supervised and unsupervised speaker adaptation. Develop state-of-the-art voice cloning framework for both text-to-speech and voice conversion.

Research Assistant

VNUHCM - University of Science - AILAB | Ho Chi Minh City, Vietnam

September 2014 - July 2017

Research and develop Automatic speech recognition and text-to-speech systems for Vietnamese language and prepare practical demonstrations to showcase the applications of these speech processing systems.

Research Intern

National Institute of Informatics - Yamagishi Lab | Tokyo, Japan

March 2016 - August 2016

Deep Learning for Statistical Parametric Speech Synthesis. DNN-based Multi-speaker speech synthesis model, adaptation and the ability to manipulate the synthetic speech characteristic.

Education

Doctor of Philosophy - Multidisciplinary Sciences, Informatics

Graduate University for Advanced Studies, SOKENDAI | Tokyo, Japan

October 2017 - September 2020

Master of Science - Computer Science

VNUHCM University of Science | Ho Chi Minh City, Vietnam

December 2014 - December 2016

Bachelor of Science - Information Technology
VNUHCM University of Science | Ho Chi Minh City, Vietnam
September 2010 - September 2014

Patents

Learning device, learning method, voice synthesis device, voice synthesis method and program
Inventors: Hieu-Thi Luong, Junichi Yamagishi
P7109071 · Issued Jul 29, 2022

Honors and Awards

Japanese Government (Monbukagakusho:MEXT) Scholarship for 3-year PhD Study
October 2017 - September 2020

Skills

Languages

- **Vietnamese** : native language
- **English** : proficient user - TOEIC: 920 (2014), IELTS-Academic: 7.0 (2015)

Technologies

- Linux, Windows, MacOS
- Shell Scripts and Commands (bash, ssh, awk, sed, …)
- Programming Languages: python, javascript, c/c++

Research and Development

- **Speech Technology**: Kaldi, ESPnet, HTK/HTS
- **Deep Learning**: pytorch, tensorflow
- **Game Development**: Unity, libGDX, pixijs
- **App Development**: android and ios
- **Web Development**: flask, nodejs, HTML/CSS

Others

- **Aspiring Artists**: pen and paper drawing/sketching, vector illustration (inkscape, etc.), design (GIMP, etc.), basic 3D modeling (Blender, etc.). (https://www.instagram.com/luonghieuthi/)

Research Publications

LaughNet: synthesizing laughter utterances from waveform silhouettes and a single laughter example
Hieu-Thi Luong, Junichi Yamagishi

October 2021, arXiv preprint arXiv:2110.04946

Preliminary study on using vector quantization latent spaces for TTS/VC systems with consistent performance
Hieu-Thi Luong, Junichi Yamagishi
June 2021, Speech Synthesis Workshop 2021 **(Refereed)**

Latent linguistic embedding for cross-lingual text-to-speech and voice conversion
Hieu-Thi Luong, Junichi Yamagishi
October 2020, Voice Conversion Challenge 2020 **(Refereed)**

Deep learning based voice cloning framework for a unified system of text-to-speech and voice conversion (Ph.D. thesis)
Hieu-Thi Luong
October 2020, preprint: http://dx.doi.org/10.13140/RG.2.2.34579.91684

NAUTILUS: a Versatile Voice Cloning System
Hieu-Thi Luong, Junichi Yamagishi
October 2020, IEEE/ACM Trans. Audio, Speech, Language Process., Vol. 28, pp. 2967-2981 **(Impact factor: 3.919)**

Bootstrapping non-parallel voice conversion from speaker-adaptive text-to-speech
Hieu-Thi Luong, Junichi Yamagishi
Dec 2019, IEEE ASRU 2019 proceedings, pp. 200-207 **(Refereed)**

Training Multi-Speaker Neural Text-to-Speech Systems using Speaker-Imbalanced Speech Corpora
Hieu-Thi Luong, Xin Wang, Junichi Yamagishi, Nobuyuki Nishizawa
Sep 2019, Interspeech 2019 proceedings, pp. 1303-1307 **(Refereed)**

A Unified Speaker Adaptation Method for Speech Synthesis using Transcribed and Untranscribed Speech with Backpropagation
Hieu-Thi Luong, Junichi Yamagishi
June 2019, arXiv preprint arXiv:1906.07414

Scaling and bias codes for modeling speaker-adaptive DNN-based speech synthesis systems
Hieu-Thi Luong, Junichi Yamagishi
Dec 2018, IEEE SLT 2018 proceedings, pp. 610-617 **(Refereed)**

Wasserstein GAN and Waveform Loss-based Acoustic Model Training for Multi-speaker Text-to-Speech Synthesis Systems Using a WaveNet Vocoder
Yi Zhao, Shinji Takaki, Hieu-Thi Luong, Junichi Yamagishi, Daisuke Saito, Nobuaki Minematsu
Sep 2018, IEEE Access, pp. 60478 - 60488 **(Impact Factor: 3.367)**

Multimodal Speech Synthesis Architecture for Unsupervised Speaker Adaptation
Hieu-Thi Luong, Junichi Yamagishi
Sep 2018, Interspeech 2018 proceedings, pp. 2494-2498 **(Refereed)**

Investigating accuracy of pitch-accent annotations in neural network-based speech synthesis and denoising effects
Hieu-Thi Luong, Xin Wang, Junichi Yamagishi, Nobuyuki Nishizawa
Sep 2018, Interspeech 2018 proceedings, pp. 37-41 **(Refereed)**

Do prosodic manual annotations matter for Japanese speech synthesis systems with WaveNet vocoder?
Hieu-Thi Luong, Xin Wang, Junichi Yamagishi, Nobuyuki Nishizawa
Mar 2018, IEICE Technical Report vol. 117, no. 517, SP2017-123, pp.215-220

Adapting and Controlling DNN-based Speech Synthesis using Input Code
Hieu-Thi Luong, Shinji Takaki, Gustav Eje Henter, Junichi Yamagishi
Mar 2017, ICASSP 2017 proceedings pp. 4905–4909 **(Refereed)**

A non-expert Kaldi recipe for Vietnamese Speech Recognition System
Hieu-Thi Luong, Hai-Quan Vu
Dec 2016, WLSI/OIAF4HLT 2016 proceedings, pp. 51-55 **(Refereed)**

A DNN-based Text-to-Speech Synthesis System using Speaker, Gender and Age Codes
Hieu-Thi Luong, Shinji Takaki, SangJin Kim, Junichi Yamagishi
Oct 2016, IEICE Technical Report vol. 116, no. 279, SP2016-48, pp. 37-42