Sungkyun **Chang**

Music Information Retrieval · Machine Learning for Audio

google.scholar | sungkyun.chang@qmul.ac.uk | mimbres.github.io/home

Education

Seoul National University — *Seoul, Korea*

M.S. in Engineering (Digital Contents Convergence Major) — *Sep. 2010 - Aug. 2013*

- Graduate School of Convergence Science and Technology
- Thesis: Modeling tonal tension in music signals
- Advisor: Dr. Kyogu Lee

B.M. (Composition Major) — *Mar. 1999 - Feb. 2008*

- College of Music
- Advisor: Prof. Donoung Lee

Research Interests

Machine Learning for Music, Speech, and Audio

Recent & Ongoing

- A generative music agent that interacts with humans through audio and symbolic prompts
- Multimodal approach to music performance assessment and feedback
- Multi-task learning approach to low-resource music transcription

Past

- Industry-scale audio content search
- Playlist prediction, coversong identification, singing voice analysis/synthesis and lyrics-to-audio alignment
- Computational music theory: tonal tension

Honors & Awards

Competition

2019	**Top 3**, WSDM Cup 2019: Spotify Sequential Skip Prediction Challenge	*Melbourne, AUS*
2018	**Winner**, MIREX 2018: Audio Cover Song Identification	*Paris, France*
2018	**Finalist**, WWW 2018 Challenge: Learning to Recognize Musical Genre	*Lyon, France*
2010	**Finalist**, Korea Telecom Econovation 1st Fair: iPhone app contest	*Seoul, Korea*
2007	**Invitation**, P.Art.y, an international digital art contest, Art Center Nabi	*Seoul, Korea*
2007	**Winner**, The 4th Computer Music Contest, Korean Electro-Acoustic Music Society	*Seoul, Korea*
2001	**Silver Award**, Samsung Electronics Software Membership Award - Multimedia	*Seoul, Korea*

Scholarship

2010-2012 **Superior Academic Performance**, Seoul National University

2010-2012 **NRF Student Research Grants**, National Research Foundation of Korea

Experience

Research Experience

Centre for Digital Music (C4DM), Queen Mary University of London — *London, UK*

Research Assistant, Collaboration with ABRSM (PI: Prof. Simon Dixon) — *Jun. 2024 - present*

- Music Performance Assessment & Feedback
  - Developing a unified model to predict marks from piano exams and generate text feedback for skill improvement.
  - Addressing key challenges of (1) multimodal representation unifying MusicXML scores, audio, and text, (2) score-to-performance alignment, (3) note mistake detection, and (4) generating feedback.
  - Researching multimodal learning, music in symbolic/audio domain, music transcription, and LLMs applications for music education.

- ImprovNet
  - Co-authored a music generative model for expressive improvisation in the symbolic domain, submitted to IJCNN 2025.
  - Proposed a model that iteratively corrupts and refines input sequences to perform controllable melodic/harmonic variations, in-filling, genre transfer, and continuation.

Research Assistant, Collaboartion with Huawei (PI: Dr. Emmanouil Benetos) *Jul. 2022 - Dec. 2023*

- Collaborated on a research project with two faculty advisors and a project manager from Huawei Europe.
- Developed a multi-instrument automatic music transcription (AMT) model specifically tailored for Pop music.
- YourMT3+
  - Released a PyTorch toolkit enabling efficient multi-task learning and cross-dataset stem augmentation in music data.
  - Addressed key challenges such as (1) the lack of data representing the wide array of styles and instruments in pop music, and (2) working with partial annotations or noisy labels of music recordings.
  - Implemented an enhanced audio encoder using Mixture of Experts and hierarchical transformer (Perceiver), and a multi-stream LM decoder to utilize partially annotated data.
  - Achieved performance improvements across 10 public datasets and pop music recordings. Presented at MLSP 2024.

Cochlear.ai *Seoul, S.Korea*

Research Scientist *Apr. 2019 - Jun. 2022*

- Neural audio fingerprint: Coupling a GPU-based nearest neighbor search method with self-supervised representation learning for music retrieval. The key aspects of this work include 1) segment-level audio identification in large-scale, 2) self-supervised contrastive learning derived from the search objective, and 3) a live audio augmentation pipeline for simulating acoustic distortions. ICASSP 2021 results outperformed conventional audio fingerprints and Google's Now-playing. From April to December in 2020, as a research lead, I completed a follow-up project with SK Telecom FLO (music streaming service) to explore practical service applications.
- SED Modeling: Sound event detection (SED) APIs are the main products of Cochlear.ai. In a team of 4 research scientists, I was responsible for improving the classification performance through implementing recent papers. The pilot study covered self-supervised learning and meta-learning for exploring applicability to SED, speaker verification and IR adaptation.

Institute for Industrial System Innovation, Seoul National University *Seoul, Korea*

Researcher, in collaboration with Kakao and Kakao Brain corp. *Jun. 2017 - Mar. 2019*

- Music Genre: Implemented an audio feature embedding model based on a variant of CNNs–dual path nets. Top 6 finalists in WWW 2018 Challenge.
- Coversong ID: Researched an improved coversong identification algorithm using pairwise cross-similarity matrices as an input to CNNs. Presented at NeurIPS Workshop, ISMIR MIREX Workshop, and ICASSP.
- Sequential Skip Prediction: Researched scalable few-shot learning algorithms for sequential skip prediction in music playlists. Various classes of metric learning and Seq2Seq architectures were compared within the real-world dataset of 1 Billion user behavior logs. Top 3 result in WSDM Spotify Challenge. Presented at WSDM Workshop.

Music and Audio Research Group (MARG), Seoul National University *Seoul, Korea*

Researcher (Project: Auto-regressive Generative Adversarial Network for Singing Synthesis and Evaluation, funded by National Research Foundation) *Jun. 2017 - Apr. 2019*

- Proposed a research road map for integration of singing voice generator and artistic performance critics neural network.
- Implemented end-to-end speech synthesis based on Wavenet, Tacotron2 and FFTnet.
- Pilot study on speech-to-singing knowledge transfer in hierarchical latent space.

Researcher (Project: Lyrics-to-audio alignment, funded by National Research Foundation) *Aug. 2015 - May. 2017*

- Researched an unsupervised learning approach to Lyrics-to-audio alignment, where the audio was assumed as mixture of singing voice and accompaniments. A basic idea was to use the patterns of vowel repetition observed in both audio and text as key features.
- Conducted a pilot study to validate the theoretical upper bound of the assumption–"using only vowels".
- Implemented a front end using unsupervised singing voice separation and voice activity detection, based on RPCA.
- Presented a two stage method: 1) using weighted-symmetric-NMF for unsupervised discovery of discriminative subspace that captures repetitive patterns in vowel acoustics, 2) spatio-temporal alignment with canonical time warping.
- Outperformed against conventional HMM+ASR-based system. Published in IEEE Access.

Researcher (Project: Onset/offset detection for singing voice, funded by Samsung and Ministry of Science, ICT & Future Planning) *Aug. 2013 - Aug. 2015*

- Researched a method searching for pairwise note onset and offset in singing voice.
- Employed Correntropy, a generalized correlation function inspired by Reyni's entropy, as a detection function.
- Proposed a simple peak picking algorithm that could simultaneously capture onset/offset from the detection function.
- Outperformed against state-of-the-art. Presented at ICASSP.

Research Assistant (Project: Lyrics-to-audio alignment, in collaboration with Naver) *Jun. 2013 - Feb. 2014*

- The whole system consisted of a singing voice enhancement, pre-trained ASR and alignment modules. My contribution was: 1) implementing alignment algorithms based on semi-supervised-NMF and DTW, 2) collecting data.

Research Assistant (Project: Smart Interactive Edutainment, funded by Seoul Business Agency) *Aug. 2011 - Aug. 2013*

- Implemented pitch detection algorithm for monophonic instruments.

Research Assistant (Project: Automatic Extraction of Rich Musical Descriptors from Musical Audio, Naver) *Aug. 2010 - Aug. 2011*

- Implemented an algorithm for instrumentation analysis, based on supervised-NMF.
- Implemented an algorithm for predicting the singer's gender. The model was composed of harmonic/percussive separation, singing voice extraction, acoustic feature extraction, and a classifier. 1 domestic patent

Center for Arts & Technologies (CATSNU), Seoul National University *Seoul, Korea*

Researcher *Feb. 2006 - Aug. 2010*

- 9 collaborative works of new media art, sound installation, and live electro-acoustics: interface design, sound design, arranging, and performance as a percussionist.
- iHEAB/Hansori: Technical assistant. A live electro-acoustics performance in Oct. 2006 at Haus der Kulturen der Welt, Berlin, Germany.
- AV Brush v2: Sound interaction design. A drawing tool as a musical instrument.
- R-monome: Clone of Monome (a matrix-type musical interface). Hardware design and software programming.
- Mixplore: Musical performance with tangible interfaces. Sound & physical computing with HMM-based gesture recognition.

Samsung Software Membership Residency *Seoul, Korea*

Undergraduate Research Member in Multimedia *May. 1999 - Oct. 2001*

- The ripple of emotion: sound programming and musical performance for interactive media artwork

Selected Professional Experience

EmotionWave *Seoul, Korea*

CMTO *Jan. 2024 - May. 2024*

- Led a team in the deployment of music generative models and transcription models for a music metaverse framework.

Mimbres *Seoul, Korea*

Founder, Real-time Audio DSP *2010 - 2011*

- Design and developed a music creation tool for mobile devices with online-game-like UIs and network play modes. The audio synthesis engine was implemented based on FMOD and open-source synthesis toolkit (STK) library.

Nexon Inc, Lycos Korea, Hyundai SEGA, Netmarble Corp. *Seoul, Korea*

Freelance Sound Designer *2008 - 2010*

- Game sound design for 11 commercial online video games, published worldwide.

10th Audio Art Festival division in Korea *Seoul, Korea & Cracow, Poland*

Co-director *2003*

- Co-directed an international experimental music festival, in collaboration with Dr. Marek Choloniewski and Academy of Music in Cracow.

Publications: Music Information Retrieval

Conference & Workshop Papers

ImprovNet: Generating Controllable Musical Improvisations with Iterative Corruption Refinement
Keshav Bhandari, Sungkyun Chang, Tongyu Lu, Fareza Enus, Louis B. Bradshaw, Dorien Herremans, Simon Colton
Submitted to International Joint Conference on Neural Networks (IJCNN), 2025

YourMT3+: Multi-instrument Music Transcription with Enhanced Transformer Architectures and Cross-dataset Stem Augmentation
Sungkyun Chang, Emmanouil Benetos, Holger Kirchhoff, Simon Dixon
IEEE International Workshop on Machine Learning for Signal Processing (MLSP), London, UK, 2024

YourMT3: a toolkit for training multi-task and multi-track music transcription model for everyone
Sungkyun Chang, Emmanouil Benetos, Simon Dixon
Digital Music Research Network One-day Workshop (DMRN+17), 2022

Neural audio fingerprint for high-specific audio retrieval based on contrastive learning
Sungkyun Chang, Donmoon Lee, Jeongsoo Park, Hyungui Lim, Kyogu Lee, Karam Ko, Yoonchang Han
2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Toronto, Canada, 2021

Sequential skip prediction with few-shot in streamed music contents
Sungkyun Chang, Seungjin Lee, Kyogu Lee
WSDM Cup Workshop, 12th ACM International Conference on Web Search and Data Mining (WSDM), Melbourne, Australia, 2019

Cover song identification using song-to-song cross-similarity matrix with convolutional neural network
Juheon Lee, Sungkyun Chang, Sang Keun Choe, Kyogu Lee
2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Calgary, Alberta, Canada, 2018

Covernet: cover song identification using cross-similarity matrix with convolutional neural network
Juheon Lee, Sungkyun Chang, Donmoon Lee, Kyogu Lee
Music Information Retrieval Evaluation eXchange (MIREX), 2018

Audio cover song identification using convolutional neural network
Sungkyun Chang, Juheon Lee, Sang Keun Choe, Kyogu Lee
Machine Learning for Audio Workshop, Neural Information Processing Systems (NeurIPS 2017), Long Beach, CA, USA, 2017

A pairwise approach to simultaneous onset/offset detection for singing voice using correntropy
Sungkyun Chang, Kyogu Lee
2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Florence, Italy, 2014

Saliency-driven model for perceptual audio onset detection
Sungkyun Chang, Kyogu Lee
The 13th International Conference on Music Perception and Cognition (ICMPC), Seoul, 4-8th August, 2014

Classification of male/female singing voice in mixed audio signals by probabilistic latent component analysis and gaussian mixture models
Sungkyun Chang, Kyogu Lee
The 2nd Conference on Pioneering Convergence Technologies, 2012

Journals

Lyrics-to-audio alignment by unsupervised discovery of repetitive patterns in vowel acoustics
Sungkyun Chang, Kyogu Lee
IEEE Access 5 (2017) pp. 16636–16648. IEEE, 2017

Thesis

Modeling Tonal Tension in Music Signals
Sungkyun Chang (Thesis Advisor: Dr. Kyogu Lee)
Master Thesis (2013). Seoul National University, 2013

Publications: New Media & Others

Conference & Workshop Papers

Offline Clustering Approach to Self-supervised Learning for Class-imbalanced Image Data
Hye-min Chang, Sungkyun Chang
arXiv preprint arXiv:2212.11444, 2022

Sonification of mood state in twitter based on ANEW analysis
Sungkyun Chang, Jaehyeuk Oh, Kyogu Lee
The 1st Conference on Pioneering Convergence Technologies, 2011

The Korean traditional music ontology (KTMO)- an ontology for the Korean traditional music linked data
Souhwan Choe, Yongtae Hwang, Sungkyun Chang, Mikyoung Kim
The 3rd International Conference on Internet (ICONI 2011), Sepang, Malaysia, December 15-19, 2011, 2011

An interface for sonification of mood state in Twitter
Jaehyeuk Oh, Sungkyun Chang, Mikyoung Kim, Kyogu Lee
HCI 2011, Korea, 2011

MixPlore: a cocktail-based media performance using tangible user interfaces
Zune Lee, Sungkyun Chang, Chang Young Lim
International Conference on Arts and Technology, Yilan, Taiwan, 2009

Journals

Classification of smartphone games based on mechanics
Yeonbi Chun, Sungkyun Chang, Tack Woo
Journal of Korea Game Society 12.6 (2012) pp. 15–24. Korea Game Society, 2012

MixPlore: a digital performance using tangible user interfaces based on cocktail mixology
Zune Lee, Sungkyun Chang, Chang Young Lim
International Journal of Arts and Technology 4.2 (2011) pp. 133–154. Inderscience Publishers, 2011

Reviewer experience

2023- **Technical Reviewer**, IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)
2023- **Technical Reviewer**, IEEE/ACM Transactions on Audio, Speech, and Language Processing
2022- **Technical Reviewer**, IEEE Signal Processing Letters
2021- **Technical Reviewer**, International Society for Music Information Retrieval (ISMIR)
2018- **Technical Reviewer**, IEEE Access

Memberships

2019- **Member**, Residency program in Campus Seoul, Google for Startups
2019- **Member**, Association for Computing Machinery
2014- **Member**, IEEE Signal Processing Society
2014- **Member**, International Society for Music Information Retrieval
2007- **Member**, Korea Electro-Acoustic Music Society

Patents

Method and apparatus for generating music fingerprints
Karam Ko, Seungjin Lee, Sungkyun Chang, Yoonchang Han, Subin Lee, Donmoon Lee, Jungsoo Park, Ilyoung Jeong, Hyungui Lim
KR Patent No.10202020113961 (2021). SK Telecom and Cochlear.ai, 2021

Apparatus and method script and scene aligning for multimedia sorting, analyzing and tagging
Sejun Kwon, Yoonchang Han, Sungkyun Chang, Kyogu Lee
KR Patent No.1020140017363 (2014). Seoul National University R&DB Foundation, 2014

Real-time musical performance feedback system for beginner musician
Sejun Kwon, Yoonchang Han, Sungkyun Chang, Kyogu Lee
KR Patent No.20150034959A (2013). Seoul National University R&DB Foundation, 2013

Skills

Programming Proficient in PyTorch/TensorFlow; Supported by TPU Research Cloud (TRC) program; Audio DSP programming with C, C++, Matlab, PD, Max/MSP, LaTeX
Music Music Theory, Music Arrangement & Production, DAWs, Live electro-acoustics, Piano, Electric bass, Perfect Pitch
Languages Korean (native), English (fluent)

References

2022- **Dr. Simon Dixon**, Professor, Centre for Digital Music, Queen Mary University of London (Current employer; Project supervisor)
2022- **Dr. Emmanouil Benetos**, Reader, Centre for Digital Music, Queen Mary University of London (Previous Employer; Project supervisor)
2010- **Dr. Kyogu Lee**, Professor, Music and Audio Research Group, Graduate School of Convergence Science and Technology, Seoul National University (Thesis supervisor; Research supervisor over the last 9 years)