

# Sungkyun Chang

MUSIC INFORMATION RETRIEVAL · MACHINE LEARNING FOR AUDIO

Cochlear.ai, B2 Autoway Tower, 06182, Seoul, KOREA

☎ (+82) 10-8801-7729 | ✉ 1ronyar@gmail.com | 📷 mimbres | 📄 scholar.google

## 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

- Self-supervised representation learning for large-scale music/audio contents analysis
- Deep audio fingerprinting

#### POTENTIAL

- Self-supervised disentanglement (e.g. style/timbre/pitch/speaker/instrument) for audio analysis and re-synthesis
- Machine listening that learns to reason by exploration and exploitation

#### PAST

- Music classification, coversong identification, onset detection for singing voice, lyrics-to-audio alignment
- Generative auto-regressive model, Non-negative matrix factorization
- Computational music theory: tonal tension

## Honors & Awards

### COMPETITION

- |      |   |                |
|------|---|----------------|
| 2019 | <b>Top 3</b> , WSDM Cup 2019: Spotify Sequential Skip Prediction Challenge            | Melbourne, AUS |
| 2018 | <b>Winner</b> , MIREX 2018: Audio Cover Song Identification                           | Paris, France  |
| 2018 | <b>Finalist</b> , WWW 2018 Challenge: Learning to Recognize Musical Genre             | Lyon, France   |
| 2010 | <b>Finalist</b> , Econovation 1st Fair: iPhone app contest                            | Seoul, Korea   |
| 2007 | <b>Invitation</b> , P.Art.y, an international digital art contest, Art Center Nabi    | Seoul, Korea   |
| 2007 | <b>Winner</b> , The 4th Computer Music Contest, Korean Electro-Acoustic Music Society | Seoul, Korea   |
| 2001 | <b>Silver Award</b> , Samsung Electronics Software Membership Award - multimedia      | Seoul, Korea   |
| 2000 | <b>Bronze Award</b> , 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

#### Cochlear.ai (start-up company)

Seoul, S.Korea

RESEARCH SCIENTIST

Apr. 2019 - PRESENT

- Neural audio fingerprint: Solo work until April 2020. My initial idea was to couple the 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 (CPC, SimCLR, Wav2Vec 2.0, BYOL, etc.) and other topics (attention, 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 in 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 in 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 - Apr. 2014*

- 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 in 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. 2008 - 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: Live electro-acoustics performance. World premiered in Oct. 2006, Haus der Kulturen der Welt, Berlin, Germany.
- AV Brush v2: A drawing tool as a musical instrument. Sound interaction design.
- 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

### Mimbres (Mobile Music Sequencer)

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 electro-acoustic music festival, in collaboration with Dr. Marek Choloniewski and Academy of Music in Cracow.

## Publications: Music Information Retrieval

### CONFERENCE & WORKSHOP PAPERS

#### 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: Sonification, Ontology, New Instrument & Game

### CONFERENCE & WORKSHOP PAPERS

#### 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

---

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.1020130114970 (2013). SEOUL NATIONAL UNIVERSITY R&DB FOUNDATION, 2013*

## Skills

---

<b>Programming</b>	Proficient in PyTorch/TensorFlow (GPU/TPU) and Matlab; Supported by TPU Research Cloud (TRC) program; Real-time Audio DSP programming with C, C++, PD, Max/MSP, STK and FMOD; 
<b>Music</b>	Harmony analysis (modern/classical), Counterpoint, Orchestration, Live electro-acoustics, Piano, Elec-bass, Perfect Pitch
<b>Languages</b>	Korean (native), English (fluent)

## References

---

- 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) E-mail: kglee@snu.ac.kr
- 2012- **Dr. Yoonchang Han**, CEO, Cochlear.ai, Seoul, Korea (Colleague for 6 years at Music and Audio Research Group; Current employer) Email: ychan@cochlear.ai