

# RONG GONG

rong.gong@upf.edu  $\diamond$  ronggong.github.io

Zhong guan cun bei yi tiao jia 5 lou 3 dan yuan 501, 100190 Beijing China

Roc Boronat, 138, 55-302, 08018 Barcelona Spain

## GOAL

---

**I am looking for a job on MIR R&D starting from February 2019.**

## EDUCATION

---

### MTG-Universitat Pompeu Fabra

May 2015-Sep 2018

PhD candidate, thesis deposited at 25th September, 2018

Barcelona, Spain

Automatic assessment of singing voice pronunciation: a case study of jingju (Beijing opera) music

<http://compmusic.upf.edu/phd-thesis-rgong>

### UPMC/IRCAM/TELECOM

September 2013-September 2014

Master 2 ATIAM

Paris, France

Acoustics, Signal Processing and Computer Science in Musical Applications

**Mention:** Awarded with honours (Assez bien)

**Thesis:** Score following for the alignment of singing voice

### Université du Maine

September 2012-July 2013

Master 1 in Acoustics and Signal processing

Le Mans, France

Acoustics, Signal Processing and Computer Science in Musical Applications

**Mention:** Awarded with honours (Assez bien) **Rank:** 1/20.

**Project:** Dimensioning of an elliptical opening on the top plate of the Chinese instrument Pi'pa

### Communication University of China

September 2008-July 2011

Master in Acoustics, Psychoacoustics and Signal Processing

Beijing, China

**Thesis:** Analysis on timbre features of the plucked-string instruments

### Northwest University

September 2004-July 2008

Bachelor in Applied Mathematics

Xi'an, China

## RESEARCH EXPERIENCES

---

### Jukedek

Nov 2018-Jan 2019

Internship

London, UK

- Develop dereverberation and de-artefacts deep learning models.

### MTG-UPF

Oct 2018

Essentia project

Barcelona, Spain

- Implement PYin (probabilistic Yin) pitch extraction algorithms in Essentia (Open-source library and tools for audio and music analysis, description and synthesis).

### MTG-UPF

April 2018-Present

TECSOME project

Barcelona, Spain

- Develop automatic solfège assessment API using KALDI.

**National Academy of Chinese Theatre Arts (NACTA)** September 2016-Present  
*Project collaborator-Beijing City Social Science Foundation 16JDYTA016* *Beijing and Barcelona*

- Collect jingju singing recording materials of both professional and amateur singers.
- Develop iOS jingju singing courseware.

**MTG-UPF** May 2015-Sep 2018  
*CompMusic project* *Barcelona, Spain*

- Develop algorithms for computer-aided jingu (Beijing opera) singing assessment system.
- Automatically retrieving Jingju singing features (syllable segmentation, singing phrase identification).

**iOS developer** October 2014-May 2015  
*iOS music application development* *Paris, France*

- Develop the iOS application - Acousticbrainz jukebox for exploring music by emotion.

**IRCAM** March 2014-August 2014  
*Master 2 internship* *Paris, France*

- Improve the performance of the score following system - Antescofo on singing voice.

**Université du Maine** March 2013-June 2013  
*Master 1 project* *Le Mans, France*

- Design the opening on the top plate of the traditional Chinese instrument - Pi'pa to augment its low-frequency performance.

**Communication University of China** September 2010-September 2011  
*Master thesis work* *Beijing, China*

- National Natural Science Foundation of China, NSFC10804100
- Analysis the timbre of traditional Chinese plucked-string instruments by musical features and subjective listening experience

**3G Audio** November 2010-January 2011  
*Master internship* *Dongguan, China*

- Study the impact of membrane materials on the sound quality of loudspeaker

## SUMMARY OF COMPETENCES

---

### Music Signal Processing

▷ Chorus detection. ▷ Melody extraction. ▷ Intonation element segmentation. ▷ Note transcription. ▷ Tempo estimation. ▷ High resolution methods for spectral estimation (MUSIC, ESPRIT).

### Machine Learning

▷ Deep learning (CNNs, RNNs). ▷ Gradient boosting machine. ▷ Data clustering (PCA, K-means, Decision tree). ▷ Bayesian inference (HMMs, HSMMs, CRFs).

### Speech Technologies

▷ Spectral subtraction (implicit Wiener filter, a priori SNR estimation Wiener filter). ▷ One step prediction Kalman filter. ▷ Noise power estimation by minimum statistics. ▷ KALDI speech recognition

## **Computer Music**

▷ Digital filter VST plugins programming. ▷ Language LISP. ▷ Constraint programming. ▷ MIDI protocol.

## **ACADEMIC ACTIVITIES**

---

Google Speech Summit traveling grant, London, May 2018

Task captain of MIREX 2018: Automatic Lyrics-to-Audio Alignment

ISMIR conference reviewer 2017, 2018

## **PUBLICATIONS IN CHRONOLOGICAL ORDER**

---

### **Rong Gong, Xavier Serra**

Singing Voice Phoneme Segmentation by Hierarchically Inferring Syllable and Phoneme Onset Positions  
Interspeech, Hyderabad, India, September 2018

### **Eduardo Fonseca, Rong Gong, Xavier Serra**

A Simple Fusion of Deep and Shallow Learning for Acoustic Scene Classification  
Sound and Music Computing Conference (SMC), Limassol, Cyprus, July 2018

### **Eduardo Fonseca, Rong Gong, Dmitry Bogdanov, Olga Slizovskaia, Emilia Gomez, Xavier Serra**

Acoustic Scene Classification by Ensembling Gradient Boosting Machine and Convolutional Neural Networks

2nd workshop on Detection and Classification of Acoustic Scenes and Events (DCASE), Munich, Germany, November, 2017

### **Rong Gong, Rafael Caro Repetto, Xavier Serra**

Creating an A Cappella Singing Audio Dataset for Automatic Jingju Singing Evaluation Research  
4th International Digital Libraries for Musicology workshop (DLfM 2017), Shanghai, China, October, 2017

### **Rong Gong, Xavier Serra**

Identification of potential Music Information Retrieval technologies for computer-aided jingju singing training

Chinese traditional music technology session - China conference on sound and music technology, Suzhou, China, October, 2017

### **Rong Gong, Jordi Pons, Xavier Serra**

Matching Audio to Score by Combining Phonetic and Duration Information

International Society for Music Information Retrieval Conference (ISMIR), Suzhou, China, October 2017

### **Jordi Pons, Rong Gong, Xavier Serra**

Score-informed Syllable Onset Detection for A Cappella Singing Voice with Convolutional Neural Networks

International Society for Music Information Retrieval Conference (ISMIR), Suzhou, China, October 2017

### **Jordi Pons, Olga Slizovskaia, Rong Gong, Xavier Serra**

Timbre Analysis of Music Audio Signals with Convolutional Neural Networks

European Signal Processing Conference (EUSIPCO), Kos island, Greece, August 2017

### **Rong Gong, Nicolas Obin, Georgi Dzhambazov, Xavier Serra**

Score-informed Syllable Segmentation for Jingju A Cappella Singing Voice with Mel-frequency Intensity

Profiles

Folk Music Analysis (FMA), Málaga, Spain, June 2017

**Rong Gong, Yile Yang, Xavier Serra**

Pitch Contour Segmentation for Computer-aided Jingju Singing Training

Sound and Music Computing Conference (SMC), Hamburg, Germany, August 2016

**Rafael Caro Repetto, Rong Gong, Nadine Kroher, Xavier Serra**

Comparison of The Singing Style Of Two Jingju Schools

International Society for Music Information Retrieval Conference (ISMIR), Málaga, Spain, October 2015

**Rong Gong, Philippe Cuvillier, Nicolas Obin, Arshia Cont**

Real-Time Audio-to-Score Alignment of Singing Voice Based on Melody and Lyric Information

Interspeech, Dresden, Germany, September 2015

**Xin Wang, Rong Gong**

Consonance of Musical Dyads of Chinese Plucked Instruments

2nd International Conference on Computer Science and Network

Technology, ICCSNT, Changchun, China, 2012

**Rong Gong, Zihou Meng**

Analysis on Tonal Components of Plucked-String Instruments Tones

National Youth Conference of Acoustical Society of China, CYCA 9th

Guangzhou, China, November 2011

**Rong Gong, Zihou Meng**

Analysis on Feature Discrimination of Plucked-string Instruments Tones

National Conference on Man-Machine Speech Communication, NCMMSC2011

Xi'an, China, October 2011

## LANGUAGES

---

<b>Mandarin</b>	Mother tongue
<b>English</b>	Diploma CLE B2
<b>French</b>	Diploma TCF B2
<b>Spanish</b>	B1

## COMPUTER SKILLS

---

<b>Programming</b>	Python, Objective-C, Swift, Matlab, C
<b>Machine learning and Deep learning packages</b>	Keras, Tensorflow, Scikit-learn
<b>Speech analysis and recognition tools</b>	KALDI, Praat
<b>Music Production</b>	Ableton Live, Cubase