Singing voice synthesis in the context of music technology research

Xavier Serra
Music Technology Group
Universitat Pompeu Fabra, Barcelona
http://mtg.upf.edu
Musical Sound Synthesis
Theremin
Lev Termen, 1917

Clara Rockmore
“Vocalise” by Rachmaninoff

RCA theremin, 1929
Moog Synthesizer
Robert Moog, 1964

Moog synthesizer from the 60s

Walter Carlos
“Switched on Bach”, 1968
FM synthesis
John Chowning, 1972

DX7 by Yamaha, 1983
Waveguide synthesis

Julius Smith, 1985

VL1 by Yamaha, 1994
Sampling

Fairlight, 1980
New musical interfaces

Reactable, 2005
Singing Voice Synthesis
Physical model from 50s

Max Mathews, 1965

Kelly, Lochbaum and Mathews, 1961
“Daisy”
Linear Predictive Coding

Subtractive Voice Synthesis Block Diagram

Noise Gen.

Voiced/Unvoiced

Amp_{0}(t)
Freq_{0}(t)

Amp
Freq

Pulse Generator at Voice Fundamental

Amp_{1}= f(t)
Freq_{1}= f(t)

Amp_{2}= f(t)
Freq_{2}= f(t)

Amp_{3}= f(t)
Freq_{3}= f(t)

Resonant (bandpass) Filters at Voice Formants

Charles Dodge
“The days ahead”, 1975
FOF synthesis
Xavier Rodet, 1977

Aria “Queen of the Night”
from the Magic Flute by Mozart, 1979
SinSy
Keiichi Tokuda et al., 2009
Vocaloid and Hatsune Miku
Database Creation

Sample editing

Harmonic + Stochastic analysis

sines
residual

Excitation + Resonance model estimation

Timbre
(Resonances, Excitation)

Stationary
Note attack
Note release
Note to note
Articulation
Vibrato
Synthesis

Flat harmonic excitation

Flat residual excitation

1/F

EpR spectral Amplitude

Filtering

EpR spectral Phase

Filtering

HpS synthesis

Demo from 2001
VOCALOID GENERATION

11.6 ONLINE

ボカロに代わって歌ってみた

You know me　っす。

Yo, what's up?

先着特典：フルカラーポストカード
付いてきます！
当ビル地下にて開催中！

夏祭 初音 鑑

月に酔うて
影に舞う
歌う花一輪
初音ミク

大好評公演中

8月21日～8月29日
ベルサール池袋

TICKETS
Now on Sale!
Hatsune Miku
Some challenges in singing voice research
Computational models
for the discovery of the World's Music

HOME

CompMusic is a research project funded by the European Research Council and coordinated by Xavier Serra from the Music Technology Group of the Universitat Pompeu Fabra in Barcelona (Spain). It aims to advance in the automatic description of music by taking a culture specific approach. It carries research within the field of information technologies but taking an interdisciplinary approach. The project focuses on five music traditions of the world: Hindustani and Carnatic music of India, makam music of Turkey, Beijing Opera, and Arab-Andalusian (Magreb) music.

Project highlights:

2nd CompMusic Workshop Program
Proceedings of 2nd CompMusic Workshop
2nd CompMusic Workshop videos
2nd CompMusic Workshop blog post

Hindustani collection
Carnatic collection
Turkish makam collection
Beijing Opera collection

CompMusic sounds

LATEST NEWS

Xavier Serra gives a keynote at SMAC/SMC 2013
25/07/2013 - 08:19

Xavier Serra gives a keynote on CompMusic at the join conference of...

CompMusic presentations at the UPF
10/07/2013 - 11:50

On July 9th we had a general meeting of the MTG in which we presented all...

Vignesh Ishwar and Kaustub Kanti Ganguli sing in Barcelona
02/06/2013 - 18:30

LATEST BLOGS

Indian Art Music concerts in Barcelona 09/07/2013
Kaustub Kanti Ganguli from IIT Bombay (India) and Vignesh Ishwar from IIT Madras (India) visited MTG for a research stay and as a part of their visit, two concerts were organized jointly by CompMusic and Phonos...

Arab-Andalusian Music 06/04/2013
On February 2013 I made a trip to Morocco, together with Mohamed Sordo, to attend an Andalusian music festival in Fez and to meet with the musician and musicologist Amin Chaouch in Tetouan. This was the first trip to Morocco in the...

Applause and Aesthetic Experience 08/01/2013
M.V.N Murthy, collaborator of CompMusic that part from being a physicist is an excellent Veena player, wrote this very nice article about the aesthetics of Indian art compared with the European one, specially talking about music.
Carnatic music

T. M. Krishna
Beijing Opera

Zhao Xiu Jun
Intonation
Melodic Motives

Discovery
Induction
Extraction

Matching
Retrieval

Transform

Characterization

N dimensions

0 200 400 600 800

W = 12 seconds

m = 3 seconds
Lyrics and music

Chinese is a tonal language

妈 ma1 — “mum”
麻 ma2 — “hemp”
马 ma3 — “horse”
骂 ma4 — “to scold”
吗 ma — “?”

Hypothesis: performers try to reflect tone contours in the melody in order to make lyrics understandable.
Conclusion

• Musical instruments <---> Voice

• Singing voice synthesis <---> Speech synthesis

• Timbre models --> Musical models

• Lessons learned from Hatsune Miku
Thanks