CFA '18 LE HAVRE ▪ 23-27 avril 2018
14ᵉᵐᵉ Congrès Français d'Acoustique

CFA18

LE HAVRE

Expressive Singing Synthesis using an Extended Source Filter Model

A. Roebel, L. Ardaillon et N. Obin

STMS- IRCAM/UPMC-Sorbonne University/CNRS, 1, place Igor-Stravinsky, 75004 Paris, France

axel.roebel@ircam.fr

We will describe Ircam's concatenative Singing Synthesis system ISiS putting the focus on the following two aspects. First, we will describe the semi-parametric speech signal model that is used to adapt pitch and duration of the recorded singing database. The model consists of an extended source filter model using a deterministic and stochastic decomposition (a pulse and noise model, PaN) including a parametric representation of the glottal source. Here we will notably discuss the underlying analysis of the glottal pulse parameters and we will discuss and demonstrate the improved signal transformation quality compared to more traditional approaches. Moreover we will discuss some recent approaches to establish high level control of the singing voice quality.

Second, we will describe the recent evolutions of the singing style model that is used to generate the control parameters (f0, intensity contour and phoneme durations). The system initially presented at CFA2016 that is used to learn singing style models from historic recordings of popular French artists has been modified to take into account dependencies between different parameters and has been extended to include the intensity contours of the singing voice.