ACOUSTICS2008/1858
F0 contour estimation based on time-varying complex speech analysis

Keiichi Funaki

Univ. of the Ryukyus, Senbaru 1, Nishihara, 903-0213 Okinawa, Japan

Robust F0 (Fundamental frequency) estimation plays an important role in speech processing such as speech coding and tonal speech recognition. We have already proposed robust F0 estimation algorithm based on time-varying complex AR (TV-CAR) speech analysis for analytic signal, in which the weighted autocorrelation function is calculated using the complex residual and then the F0 is searched as the peak sample for each frame [IEICE Trans. on Fundamentals, Vol.E91-A.No.3][IEICE Trans. on Fundamentals, Vol.E90-A, No.8]. Although the algorithm can estimate more robust F0 estimation for IRS filtered speech corrupted by additive noise, the algorithm cannot perform better for non-IRS filtered speech or slightly contaminated IRS-filtered speech. In addition, the frame-based F0 estimation cannot extract the F0 trajectories in the time-domain. In order to cope with the drawbacks, this paper proposes quite simple F0 contour estimation algorithm based on the TV-CAR speech analysis, in which the F0 contour is estimated by peak-picking for the estimated time-varying spectrum that is the same manner as formant frequency estimation. The experimental results demonstrate that the proposed method leads to more accurate continuous F0 estimation than the conventional one for high-pitched speech due to the nature of analytic signal for non-IRS filtered high-pitched female speech.