

ACOUSTICS2008/935

A temporally stable representation of power spectra of periodic signals and its application to F0 and periodicity estimation

Hideki Kawahara^a, Masanori Morise^a, Toru Takahashi^a, Ryuichi Nisimura^a, Hideki Banno^b and Toshio Irino^a

^aFaculty of Systems Engineering, Wakayama University, 930 Sakaedani, 640-8510 Wakayama, Japan

^bMeijyo University, 501-1 Shiogamaguchi, Tenpakuku, 468-8502 Nagoya, Japan

A unified framework for extracting F0 and periodicity is proposed based on a new formulation of STRAIGHT. Averaging power spectra, they are calculated using two time windows a half pitch period apart, yields a temporally stable power spectrum for periodic signals (TANDEM spectrum). Applying consistent sampling theory also yields computationally efficient implementation of spectral envelope estimation used in STRAIGHT (STRAIGHT spectrum). Normalized TANDEM spectrum by STRAIGHT spectrum only represents periodicity information and can be used to detect periodicity both in the base-band (for F0 extraction) and other frequency bands (for mixed mode excitation). This unified framework provides a basis to represent the complex modes of voice excitation and enables greatly realistic remaking of extreme vocal expressions. [Supported by Grant in aids for scientific research and CrestMuse project]