PITCH DETERMINATION OF SPEECH SIGNALS BY NONLINEAR DIGITAL FILTERING

Wolfgang Hess, L.f. Datenverarbeitung, TU München, W. Germany

Pitch determination can be done in many ways. In the time domain, the first harmonic can be enhanced by low-pass filtering, or the temporal structure of the signal can be changed in such a way that periodicity is easily detected. Pitch detectors of this type, however, get into trouble when the first harmonic is attenuated or missing. To overcome this problem, the first harmonic must be reconstructed by nonlinear distortion. To study this effect, several nonlinear functions (NLFs) were examined in order to select one that can be applied to any signal within the whole range of pitch. No function, however, meets this requirement with optimal performance. Thus a combination of three NLFs (odd, even, and SSB) was selected, giving a good approximation to the ideal case. Using these NLFs, a given pitch detector (Hess, 1976) has been modified so as to make it independent of the type of input signal. The signal is first simultaneously processed by the three NLFs. The subsequent linear filtering steps represent a crude approximation to the inverse filters. A low-pass filter removes the higher formants (separately for each of the 3 channels). Then F1 or, for the even NLF, the dominating frequency resulting from the filtering after distortion is determined in each channel. The subsequent adaptive band-stop filter removes this dominating frequency; its zero, however, is commonly adjusted for the 3 channels to the highest of the "formant" frequencies actually measured. This ensures that the first harmonic is never suppressed, even when it coincides with F1. Hence, the output signal of the band-stop filter contains a strong first harmonic at least in one channel. Deriving preliminary pitch period boundaries (markers) in each channel, and checking the regularity of these markers during short intervals (25 ms), the algorithm selects the appropriate channel for final processing. In a preliminary test, the algorithm showed good performance for undistorted as well as for band-limited signals within a range of fundamental frequencies from 70 to 500 Hz.

Reference

Hess, W. (1976): "An algorithm for time-domain pitch period determination", IEEE Intern. Conf. Acoust., Speech, and Signal Processing, Philadelphia PA (ICASSP-76), 322-325.