International Journal of Innovative
Computing, Information and Control
Volume **5**, Number **9**, September **2009**

 ISSN 1349-4198
pp. 2717–2724

PITCH DETECTION USING CIRCULAR AVERAGE MAGNITUDE DIFFERENCE FUNCTION BASED ON WAVELET TRANSFORM

Jing Dong[1], Xiaopeng Wei[1,*], Qiang Zhang[1,*] and Yanping Zhao[2]

[1]Liaoning Key Lab of Intelligent Information Processing
Dalian University
Dalian, 116622, P. R. China
*Corresponding author: { xpwei; zhangq }@dlu.edu.cn

[2]College of Communication Engineering
Jilin University
Changchun, 130012, P. R. China

Received February 2008; revised July 2008

Abstract. *Robust pitch detection is one of the most important technologies in speech signal processing. According to the transmissibility of signal discontinuity under different resolution of wavelet transform, and the appearance of valley values at multiples of pitch period in the average magnitude difference function (AMDF), a new method for pitch detection based on the above is presented in this paper. In this method, a wavelet transform is used for enhancing the periodicity of noisy speech, then, the pitch is estimated by locating the most suitable valley value in the circular average magnitude difference function of the low frequency coefficients. Simulation results indicate that the proposed algorithm possesses lower calculation complexity, better pitch detection precision under strong background noise, and the capability for real time implementation.*
Keywords: Pitch detection, Wavelet transform, Circular average magnitude difference function

1. **Introduction.** The quasi-periodic vibration of vocal fold produces the excitation of voiced speech, and pitch period presents the periodicity of vocal fold vibrations at the glottis. It is one basilic speech parameter and essential to a variety of speech processing applications. For this reason, many methods have been proposed in this area, but, in a reallife situation, they must face the challenge in presence of noise. For example, the performances of comb filter speech enhancement and speaker recognition system are closely related to the accuracy of pitch period detection in practical noisy environments. Unfortunately, most methods are sensitive to noise, not reliable for wide range of pitch periods and robust pitch detection is still one intensively studied problem.

Using low-frequency band of noisy speech, C. Shahnaz first formulated an effective noise reduction approach for power spectral subtraction to track the time-variation of the non-stationary noise, then used normalized circular difference function of the enhanced speech to extract pitch [1]. Although the method performs well, the most contribution lies in the anterior speech enhancement procedure, not in the robust pitch extraction. Among the conventional methods, autocorrelation-based and AMDF-based are the most traditional algorithms applied in many systems [2, 3, 4]. T. Shimamura proposed weighted autocorrelation (WAC) algorithm, the autocorrelation function was weighted by the reciprocal of the AMDF [2]. The weighted feature can further emphasize the curve peak and improve the accuracy of pitch extraction for clean speech. But due to the falling trend of the notches in AMDF of noisy signal, which often leads to half and double pitch errors, some non-pitch peaks were emphasized and misguidances occurred. W. Y. Zhang