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

ICIC International ©2009 ISSN 1349-4198
pp. 2725–2733

SPEECH ENHANCEMENT ALGORITHM BASED ON HIGHER-ORDER CUMULANTS PARAMETER ESTIMATION

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

[1]School of Mechanical Engineering
Dalian University of Technology
Dalian, 116024, P. R. China
*Corresponding author: xpwei@dlu.edu.cn

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

Received February 2008; revised July 2008

Abstract. *A robust and effective speech enhancement algorithm is proposed for additive Gaussian noisy speech signal. When applied to speech enhancement, the fundamental difference between the traditional extended Kalman filter and our method is that we obtain the AR parameter estimations by using conjugate gradient algorithm to solve the modified Yule-Walker equation of higher-order cumulants. Simulation results show that this algorithm possesses high parameter estimation accuracy and robustness. At the same time, it can obtain good speech enhancement performance in the presence of very complicated noise.*

Keywords: Speech enhancement, Parameter estimation, Higher-order cumulants, Kalman filter, Conjugate gradient (CG)

1. Introduction.

During the mobile communication, speech signal is often affected and influenced by physical environment noises, which exerts a negative influence on speech quality. However, most speech processing systems, speech coding and automatic speech recognition system, are designed for clean speech and are relatively easy to accomplish fairly complex tasks in controlled quiet laboratory environments. When used in a real-life situation, the performance of the system deteriorates severely, which is the most major obstacle to the commercial use of speech processing technology. Therefore, as an important preconditioner, speech enhancement aims at extracting clean speech under every kind of background noise, while minimizing distortion of speech and keeping residual noise sounding natural. It is significant and necessary in communication and many speech signal processing applications.

With the development of signal processing theories and techniques, in recent years, many effective algorithms have been developed in this area for different background noise. The most popular methods are based on spectral subtraction algorithm since they are attractive for reducing additive stationary noise in a simple and efficient manner, but they usually result in musical residual noise annoying to human ear [1]. Beam forming is an emerging technique and can realize speech enhancement under non-stationary condition, but it often needs many microphones [2, 3, 4]. Moreover, many current algorithms still present important limitations, particularly due to the fact that they only focus on one given noise. As a result, the techniques are becoming more and more complex with noise diversification, and handling various kinds of noise remains a challenging task. At